BULLETIN

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

EARLE R. HEDRICK

DAVID R. CURTISS WILLIAM R. LONGLEY

WITH THE ASSISTANCE OF

HAROLD T. DAVIS HARRY W. KUHN

THEOPHIL H. HILDEBRANDT DERRICK N. LEHMER

JOHN R. KLINE HENRY L. RIETZ

CAROLINE E. SEELY

VOLUME XXXVII, NUMBER 4

APRIL, 1931

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1931

PUBLISHED MONTHLY

Whole No. 379 $9.00 a Year

Entered as second class matter at the post office at Menasha, Wis., under the act of August 24, 1912.

PUBLICATIONS OF THIS SOCIETY

I. THE TRANSACTIONS. To members, $6.75; to others, $9.00.
Volumes 1-14 and 16-21: to members, $5.00; to others, $7.00.
Volumes 15, and 22-32: to members, $6.00; to others, $8.00.

II. THE BULLETIN. To non-members, $9.00.
Members of the Society receive the Bulletin if they have paid their dues, $5.00 of which is for subscription to the Bulletin.
Volumes 1-11 (and the three volumes of the Bulletin of the New York Mathematical Society), are out of print. The Society is glad to receive odd volumes or even odd numbers of these volumes.
Volumes 12-36: to members, $5.25; to others, $7.00.

III. THE AMERICAN JOURNAL OF MATHEMATICS
Published by this Society and the Johns Hopkins University. Send orders to The Johns Hopkins Press.
To members, $5.75; to others, $7.50.
Complete sets (50 volumes): $350.00.

IV. THE COLLOQUIUM PUBLICATIONS.
Published in December, 1930:
S. Lefschetz, *Topology*, $4.50. (Volume XII of the Colloquium Series.)
To appear early in 1931:
O. Veblen, *Analysis Situs*. Second edition. $2.00. (Part II of Volume V of the Colloquium Series.)
For earlier issues see a fly-leaf of the March issue.

Orders may be addressed to
450-458 AHNAIP ST., MENASHA, WIS., OR
501 WEST 116TH STREET, NEW YORK, N.Y.

On sale also by the following official agents of this Society:
Bowes & Bowes, 1 Trinity St., Cambridge, England.
Hirschwaldsche Buchhandlung, Unter den Linden, 68, Berlin.
Librairie Scientifique Albert Blanchard, 3 bis, Place de la Sorbonne, Paris, V, France.
Libreria Editrice Nicola Zanichelli, Bologna, Italy.

LINEAR FUNCTIONAL TRANSFORMATIONS IN GENERAL SPACES†

BY T. H. HILDEBRANDT

1. *Introduction.* An abstract theory of linear functional transformations has as guide linear transformations in a finite or denumerably infinite set of variables, linear integral transformations and equations associated with these. The desire to proceed symbolically and replace details by general procedure seems to be inherent in the situation. Pincherle‡ is perhaps one of the first great exponents, so that he even seems to have anticipated some of the famous results of integral equations by a number of years. E. H. Moore§ set himself the task of unifying the Fredholm theory of integral equations and algebraic equations in finitely and infinitely many variables, and has succeeded in setting up a system which indicates in a host of special cases a valid and elegant method of procedure analogous to the Fredholm integral equation theory. Volterra¶ has devised an elegant theory of linear integral and associated operations based on the notion of permutability or commutativity of operations.

The theory to which the main portion of this address is devoted has not been, in the main, presented as such in published form. It is, however, obvious that F. Riesz in his book entitled *Les Equations Linéaires à une Infinité d'Inconnus*‖ and his paper, *Lineare Funktionalgleichungen*†† has had in mind the generalizations treated here. As a consequence, what is given here is in the main hardly new, excepting that by presenting it from the point of view of a general basis, there is a gain in elegance and simplicity for the non-pathological results.

† An address delivered at the Summer Meeting of the Society at Providence, September 10, 1930, by invitation of the program committee.

‡ See *Notice sur les travaux*, Acta Mathematica, vol. 46 (1925), pp. 341–362, especially p. 347, and pp. 351–354; *L'Operazioni Distributive*, Bologna, 1901; Encyclopédie des Sciences Mathématiques, vol. II, 26 (1912).

§ See this Bulletin, vol. 28 (1912), pp. 334–362; Proceedings of the Cambridge International Congress, 1912, vol. I, pp. 230–255.

¶ See for instance *Leçons sur les Fonctions des Lignes*, Paris, 1913, Chap. 9, etc.

‖ Paris, 1913; this book will be cited in what follows as Riesz, *Inf. Inc.*

†† Acta Mathematica, vol. 41 (1917), pp. 75–88; this paper will be cited in what follows as Riesz, *LFG.*

The first five sections of this paper are devoted to a presentation of the theory of the linear limited transformation, indicating results which can be derived in a complete vector space. The last sections give a brief account of the generalization of the Hilbert theory of symmetric kernels, and infinite limited matrices, on a more special type of linear vector space. A general basis in which the fundamental form is obtained constructively was derived by E. H. Moore in his second theory† which aimed at a generalization of the Hilbert-Hellinger theory. The postulational basis used here was given by von Neumann.‡

2. *The Space.* We shall assume as basis a class of elements ξ constituting a *linear vector space*§ $\mathfrak{S}$, that is, a space satisfying the following conditions:

(A) There is defined addition, commutative and associative with respect to the elements; and multiplication by complex numbers, commutative, associative, and distributive; further a unique zero element 0, identical with $0\cdot\xi$ for every ξ of the space.

(B) There exists a *norm* or *metric* transforming each element into a positive real number, satisfying the conditions

(1) $\|\xi_1+\xi_2\| \leqq \|\xi_1\| + \|\xi_2\|$;

(2) $\|a\xi\| = |a| \cdot \|\xi\|$, for every complex number a;

(3) $\|\xi\| = 0$, if and only if $\xi = 0$.

(C) The space is *complete*, that is, if the sequence ξ_n satisfies the condition $\lim_{n,m}\|\xi_n-\xi_m\| = 0$, then there exists a ξ such that $\lim_n\|\xi_n-\xi\| = 0$. If the space $\mathfrak{S}$ is not complete, it is obvious that completeness is extensionally attainable after the manner of deriving the real number system from the rational number system.

As examples of spaces satisfying the above postulates, we might note the following:¶

† Available at present only in lecture notes, excepting some indications in Mathematische Annalen, vol. 86 (1922), pp. 34, etc.

‡ Göttinger Nachrichten, 1927.

§ See Banach, Fundamenta Mathematicae, vol. 3 (1922), p. 135; Wiener, Bulletin de la Société Mathématique de France, vol. 49 (1921), pp. 123–124; Fréchet, *Les Espaces Abstraits*, pp. 125–126.

¶ This list is intended to be illustrative, not exhaustive. Hahn, Monatshefte, vol. 32 (1922), pp. 1–88, gives a large number of illustrations.

GROUP I. Let ξ be a function on the set of all positive integers, that is, a sequence $\{x_p\}$, $(p=1, 2, 3, \cdots)$.

I.1$_e$. The space $\mathfrak{S}$ consists of all sequences for which $\sum |x_p|^{1+e}$ converges, the norm $\|\xi\|$ being the $(1+e)$th root of this sum. For $e=1$, we get Hilbert space.

I.2. The space $\mathfrak{S}$ consists of all convergent series, with $\|\xi\| = LUB \left|\sum_1^n x_p\right|$, where LUB is the least upper bound.

I.3. The space $\mathfrak{S}$ consists of all sequences having a limit, with $\|\xi\| = LUB\, |x_p|$.

I.3$_0$. The space $\mathfrak{S}$ consists of all sequences having zero as limit, with same norm as in I.3. Obviously there exists a one to one correspondence between I.2, 3, and 3$_0$.

I.4. The space $\mathfrak{S}$ consists of all bounded sequences with norm $\|\xi\| = LUB\, |x_p|$.

GROUP II. Let ξ be a real or complex valued function $x(p)$ on the linear interval $0 \leqq p \leqq 1$.

II.1$_e$. The space $\mathfrak{S}$ consists of all measurable functions for which the Lebesgue integral $\int_0^1 |x(p)|^{1+e} dp$ exists, with $\|\xi\|$ the $(1+e)$th root of this integral.

II.2. The space $\mathfrak{S}$ consists of all measurable functions, bounded except for a set of zero measure, $\|\xi\|$ being the effective upper bound of $|x(p)|$, that is, the greatest lower bound of the numbers M such that the set of elements p for which $|x(p)| > M$ is of measure zero.

II.3. The space $\mathfrak{S}$ consists of all continuous functions, $\|\xi\|$ being the maximum of $|x(p)|$.

II.4. The space $\mathfrak{S}$ consists of all functions of bounded variation, $\|\xi\|$ being the total variation on the interval.

For two interesting linear spaces, it is not known that they are of this vector type; these are the set of all sequences, with $\lim \xi_n = \xi$ defined as $\lim x_n(p) = x(p)$ for every p, and the set of all measurable functions with limit defined as convergence in the measure, that is, for every e, the measure of the set of elements p for which $|x_n(p) - x(p)| > e$ approaches zero with n. In each of these cases there exists a metric† giving an equivalent definition of limit but the metric does not satisfy the multiplicative condition (2).

† See, for instance, Fréchet, Rendiconti di Palermo, vol. 22 (1906), p. 38; and Bulletin of Calcutta Mathematical Society, vol. 11 (1921), pp. 190, etc.

As is customary, we shall call a subset $\mathfrak{S}_0$ of the space $\mathfrak{S}$ linear if every linear combination of elements of $\mathfrak{S}_0$ belongs to $\mathfrak{S}_0$; closed if the limit of a sequence of elements of $\mathfrak{S}_0$ also belongs to $\mathfrak{S}_0$. Of importance are the linear and closed extensions of a subset $\mathfrak{S}_0$, attainable by adding to $\mathfrak{S}$ all linear combinations of its elements, that is, the linear extension of $\mathfrak{S}_0$, and then the limits of all convergent sequences of this latter class. If $\mathfrak{S}_0$ is the linear closed extension of a finite number of elements of $\mathfrak{S}$, it is said to be of finite dimension. An important theorem in this connection is due to Riesz† that *a necessary and sufficient condition that a linear closed subset $\mathfrak{S}_0$ of $\mathfrak{S}$ be of finite dimension is that every bounded set of elements be compact* (that is, contain a subsequence having a limit).

Next to the sets of finite dimension are the sets or spaces of denumerably infinite dimension, that is, in which there exists a sequence of elements $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ whose linear closed extension is the space. The spaces I.1$_c$, I.2, I.3, II.1$_c$, II.3 have this property, but not the spaces I.4, II.2 and II.4.

A set $\mathfrak{G}$ is called a *fundamental set*, if $\mathfrak{S}$ is the linear closed extension of $\mathfrak{G}$.

3. *Linear Limited Operations.* A *linear limited operation* $L(\xi)$ on $\mathfrak{S}$ satisfies the following three conditions:

(a) To every ξ of $\mathfrak{S}$ there corresponds a real or complex number.

(b) If ξ_1 and ξ_2 are elements of $\mathfrak{S}$ and a_1 and a_2 complex numbers, then

$$L(a_1\xi_1 + a_2\xi_2) = a_1L(\xi_1) + a_2L(\xi_2).$$

(c) There exists a number M such that for every ξ of $\mathfrak{S}$, we have $|L(\xi)| \leq M\|\xi\|$. The greatest lower bound of all possible M's satisfying this condition will be called the *modulus* M_L of L. The condition (c) is equivalent in our space $\mathfrak{S}$ to continuity, that is, if $\lim_n \xi_n = \xi$, then $\lim_n L(\xi_n) = L(\xi)$, continuity at one point being sufficient.‡ In the case of metric spaces not satisfying the multiplicative condition (2) on the norm, this equivalence is not possible, and it seems preferable to discuss linear continuous operations.

The question, what is the most general linear limited operation on a given space $\mathfrak{S}$, is of interest, but has not yet received

† Riesz, *LFG*, pp. 77–79.

‡ See Banach, loc. cit., p. 152.

exhaustive treatment. The following results are comparatively well known:†

I.1$_e$, with $e>0$. $L(\xi)=\sum a_p x_p$, where $\sum |a_p|^{1+1/e}<\infty$, that is, a_p belongs to I.1$_{1/e}$.

I.1$_0$, that is, $e=0$. $L(\xi)=\sum a_p x_p$, where the set a_p is bounded, that is, a belongs to space I.4.

I.2 $L(\xi)=\sum a_p x_p$, with $\sum |a_p-a_{p+1}|<\infty$.

I.3$_0$ $L(\xi)=\sum a_p x_p$, with $\sum |a_p|<\infty$, that is, a_p is of I.1$_0$.

I.3 $L(\xi)=\sum a_p x_p+a_0 x_0$, where $\lim_p x_p=x_0$ and $\sum |a_p|<\infty$.

II.1$_e$ ($e>0$.) $L(\xi)=\int \alpha(p)x(p)dp$, where $\int |\alpha(p)|^{1+1/e}$ exists.

II.1$_0$ ($e=0$.) Same form, but α is bounded except for a set of zero measure, that is, of II.2.

II.3 $L(\xi)=\int x(p)d\alpha(p)$, α of bounded variation.

The most general linear limited operations in the spaces I.4, II.2 and II.4 do not seem to have been determined, the problem in the cases listed being simplified because of the separability of the space. Examples of linear operations in II.4 are $\int \alpha(p)dx(p)$, where $\alpha(p)$ is continuous, or more generally if a Lebesgue-Stieltjes integral, then $\alpha(p)$ may be a Borel measurable function. On the other hand $\sum_i [x(p_i+0)-x(p_i-0)]$, where p_i are the points of discontinuity of x and the summation is extended over these, is obviously linear limited, but not expressible in integral form.

The following theorem is interesting in the theory of linear operations.

THEOREM 1. *If $L_n(\xi)$ is a sequence of linear limited operations converging to the operation $L(\xi)$, then the sequence is uniformly limited* (that is, the sequence of moduli M_{L_n} is bounded) *and $L(\xi)$ is a linear limited operation.*‡

It is possible to generalize this theorem in a number of ways. First the linearity condition on L_n may be replaced by a property similar to the triangle property of the norm, namely,

† See Riesz, Mathematische Annalen, vol. 69 (1910), p. 475; Annales de l'École Normale, (3), vol. 31 (1914), pp. 9–14; Helly, Monatshefte, vol. 31 (1921), p. 84; Hahn, loc. cit.

‡ See Banach, loc. cit., p. 157.

$$| L(a_1\xi_1 + a_2\xi_2) | \leqq | a_1 | | L(\xi_1) | + | a_2 | | L(\xi_2) |;$$

for all ξ_1 and ξ_2 of $\mathfrak{S}$ and all complex numbers a_1 and a_2. Then the condition that the sequence of values $L_n(\xi)$ converge for every ξ may be replaced by assuming that for each ξ the sequence of numbers $L_n(\xi)$ is bounded. Finally the set $1, 2, \cdots, n, \cdots$ may be replaced by a general set of elements q, in which there is a transitive order relation, and any two elements have a common successor. However the boundedness of $L_q(\xi)$ has a certain uniform character, in that there exists a sequence $\{q_n\}$ independent of ξ such that for each ξ the $L_q(\xi)$ are bounded after a certain q_n.† Then the set $L_q(\xi)$ is ultimately uniformly limited.‡

The definition of bilinear and n-linear limited operations is almost self-evident, a form $B(\xi_1, \xi_2)$ being bilinear on the classes $\mathfrak{S}_1$ and $\mathfrak{S}_2$, if it transforms every element (ξ_1, ξ_2) of $(\mathfrak{S}_1, \mathfrak{S}_2)$ into a real number, and is linear on $\mathfrak{S}_1$ and $\mathfrak{S}_2$ separately; it is limited if there exists an M such that for every ξ_1 and ξ_2

$$| B(\xi_1, \xi_2) | \leqq M \| \xi_1 \| \cdot \| \xi_2 \| .$$

The spaces $\mathfrak{S}_1$ and $\mathfrak{S}_2$ do not need to coincide, and consequently the symbol $\| \ \|$ may have different interpretations.

The question of the general form of a limited bilinear operation on two given spaces has not received very much consideration, and is somewhat beset with difficulty.§ A slight inroad can be made by noting that if we fix the element ξ_2 of $\mathfrak{S}_2$, then $B(\xi_1, \xi_2)$ is a linear limited operation on $\mathfrak{S}_1$ and conversely. This gives results for the cases of Group I, where the elements ξ are sequences, and where the general linear limited operation is known. For instance, if $\mathfrak{S}_1 = \mathfrak{S}_2 = \mathrm{I}.1_1 =$ Hilbert space, then

$$B(\xi_1, \xi_2) = \sum_p \sum_q x_p^{(1)} a_{pq} x_q^{(2)} = \sum_q \sum_p x_p^{(1)} a_{pq} x_q^{(2)},$$

complete conditions on a_{pq} in addition to $\sum_p |a_{pq}|^2 < \infty$ and $\sum_q |a_{pq}|^2 < \infty$ which make B limited being still undetermined.¶

† This has an analogue in the sequential limit. The need of such a condition appears from examples given by von Neumann, Mathematische Annalen, vol. 102 (1929), p. 380.

‡ See Hildebrandt, this Bulletin, vol. 29 (1923), p. 311.

§ But see Fréchet, Transactions of this Society, vol. 16 (1915), pp. 216–234; and Radon, Wiener Berichte, vol. 122 (IIa) (1913), pp. 1381, etc.

¶ See Hellinger and Toeplitz, Encyklopädie der Mathematischen Wissenschaften, vol. IIc 13, p. 1426.

If $\mathfrak{S}_1 = \mathfrak{S}_2 = \mathrm{I}.1_e$ with $e = 0$, then B has the same form as above, the a_{pq} being subject to the condition of being uniformly bounded ($\leqq M$, the modulus of the operation). The situation is more complicated in the cases of Group II or a combination of the two groups.

By a repeated application of the generalized theorem on sets of linear limited operations, we can get a corresponding theorem on sets of bilinear limited operations, the well known theorem of Toeplitz† which states that if the bilinear form $\sum_p \sum_q x_p a_{pq} y_q$ converges for every ξ and η of Hilbert space for which $\|\xi\| = \|\eta\| = 1$, then it is limited, being a corollary of this theorem by a consideration of the bilinear operations

$$B_{mn}(\xi, \eta) = \sum_{p=1}^{m} \sum_{q=1}^{n} x_p a_{pq} y_q .$$

4. *Linear Limited Transformations.* A transformation T will be called a *linear limited transformation* on $\mathfrak{S}_1$ to $\mathfrak{S}_2$ if it satisfies the following conditions:

(a) To every element ξ of $\mathfrak{S}_1$ there corresponds an element η of $\mathfrak{S}_2$.

(b) For every ξ_1 and ξ_2 of $\mathfrak{S}_1$ and all complex numbers a_1 and a_2

$$T(a_1\xi_1 + a_2\xi_2) = a_1 T(\xi_1) + a_2 T(\xi_2) .$$

(c) There exists a number M such that for all ξ of $\mathfrak{S}_1$

$$\|T(\xi)\| \leqq M\|\xi\| .$$

The smallest possible value of M in (c) will be called the *modulus* ($M(T)$) of T. Obviously the norm for $\mathfrak{S}_1$ and $\mathfrak{S}_2$ need not be the same. As in the case of operations, condition (c) is equivalent to continuity. Also continuity at a single point is sufficient, with linearity (b), to guarantee continuity at all points and hence limitedness.

This concept covers a rather wide variety of notions, because of the generality of $\mathfrak{S}_1$ and $\mathfrak{S}_2$. For instance, a linear limited operation is a transformation, the space $\mathfrak{S}_2$ being the set of all complex numbers. A sequence of linear limited operations on $\mathfrak{S}$ converging for each member of $\mathfrak{S}$ can be considered a linear

† Mathematische Annalen, vol. 69 (1910), pp. 321–322.

limited transformation on $\mathfrak{S}$ to space I.3, which classes certain types of summability definitions as special instances of linear limited transformations.†

If the transformation T is not defined on the entire space $\mathfrak{S}$, but only on a fundamental set $\mathfrak{G}$, then it can be extended to a linear limited transformation on $\mathfrak{S}$, if, and only if, there exists an M such that for every finite subset $\xi_1, \cdots, \xi_n$ of $\mathfrak{G}$, and every set of numbers $a_1, \cdots, a_n$, we have‡

$$\left\| \sum_{i=1}^{n} a_i T(\xi_i) \right\| \leqq M \left\| \sum_{i=1}^{n} a_i \xi_i \right\| .$$

The question what form a linear limited transformation must take, is in some cases rather difficult to answer. The best chance for success is in the case in which the space $\mathfrak{S}_2$ is a set of functions, in which the existence of a norm has as consequence the boundedness of the functional values. In such cases, the values of $\eta(p)$, for p fixed, form a linear limited operation on $\mathfrak{S}_1$. For example, if $\mathfrak{S}_2$ is any of the classes of Group I, the values $\eta(p)$ are certain sequences of linear limited operations. (a) On $\mathrm{I.1}_1$ to $\mathrm{I.1}_1$

$$y_p = \sum_q a_{pq} x_q ,$$

a_{pq} being coefficients of a limited bilinear form. (b) On $\mathrm{I.1}_0$ to $\mathrm{I.1}_1$ the transformation takes the same form, but the conditions on a_{pq} are that the sums $\sum_q |a_{pq}|^2$ shall be uniformly bounded. (c) On $\mathrm{II.1}_1$ to $\mathrm{I.1}_1$, a T is of the form $\int f_p(q)x(q)$, where the f_p are of Lebesgue integrable square, and such that for every $x(q)$ of the same class we have $\sum_p |\int f_p x|^2 < \infty$. What this requires of the sequence f_p has been developed by A. J. Pell.§ A sufficient but not necessary condition is that the f_p form a normed orthogonal system. (d) On $\mathrm{II.1}_1$ to I.4 a T is of the form as in (c), where the integrals $\int f_p x$ form a uniformly bounded sequence for each ξ of II.1. For this it is necessary and sufficient that $\int |f_p|^2$ form a bounded sequence. (e) A T on II.3 to II.3 is of the form $\int x(p)d\alpha(p, q)$, where α is of bounded variation for

† See Schur, Journal für Mathematik, vol. 151 (1920), p. 79; Hahn, loc. cit., has given extended consideration of the conditions to be satisfied by special forms of operations so as to be transformations of the type mentioned.

‡ See Hahn, Journal für Mathematik, vol. 157 (1927), p. 216.

§ Transactions of this Society, vol. 12 (1911), p. 142.

every q, as a matter of fact must be uniformly of bounded variation in p. In addition it must satisfy certain continuity conditions. The case when T is on $\mathrm{II}.1_1$ to $\mathrm{II}.1_1$ is most easily expressed in terms of the indefinite integrals of the functions involved and Hellinger integrals.†

It is possible to set up an algebra of linear limited transformations. For instance, if T_1 and T_2 are on $\mathfrak{S}_1$ to $\mathfrak{S}_2$, then $c_1T_1+c_2T_2$ is defined as $T_1(c_1\xi_1)+T_2(c_2\xi_2)$. Obviously $M(c_1T_1+c_2T_2) \leqq |c_1|M(T_1)+|c_2|M(T_2)$. If we consider the totality of all limited linear transformations on $\mathfrak{S}_1$ to $\mathfrak{S}_2$, then this can be considered as a vector space of the same type as $\mathfrak{S}$, the norm being the modulus $M(T)$.

If T_1 is a linear limited transformation on $\mathfrak{S}_1$ to $\mathfrak{S}_2$, and T_2 on $\mathfrak{S}_2$ to $\mathfrak{S}_3$, then the succession T_1T_2 will define a linear limited transformation on $\mathfrak{S}_1$ to $\mathfrak{S}_3$, and it is obvious that the modulus $M(T_1T_2)$ satisfies the condition

$$M(T_1T_2) \leqq M(T_1)M(T_2).$$

Obviously T_1 and T_2, even when $\mathfrak{S}_1=\mathfrak{S}_2=\mathfrak{S}_3$, are not necessarily commutative. It follows from the definitions and linearity that $T_1(T_2+T_2')=T_1T_2+T_1T_2'$, and that three operations T_1, T_2, T_3 are associative, that is, $(T_1T_2)T_3=T_1(T_2T_3)$.

Of prime interest is the study of reciprocal relations, that is, answering the questions (a) under what conditions does there exist, for a given η of $\mathfrak{S}_2$, a ξ of $\mathfrak{S}_1$ such that $T\xi=\eta$; and more generally (b) when is this possible for every η of class $\mathfrak{S}_2$? The latter is closely allied with the question under what conditions does there exist for a given T on $\mathfrak{S}_1$ to $\mathfrak{S}_2$ a linear limited transformation T^{-1} on $\mathfrak{S}_2$ to $\mathfrak{S}_1$ such that $T^{-1}T=I$ and $TT^{-1}=I$, the identity transformation I being on $\mathfrak{S}_1$ to $\mathfrak{S}_1$ in the first instance, and on $\mathfrak{S}_2$ to $\mathfrak{S}_2$ in the second. If the first of these equations is satisfied we call T^{-1} a left hand reciprocal and if the second, then T^{-1} is a right hand reciprocal. It is obvious that if T has a right hand reciprocal T_r^{-1}, then $\xi=T_l^{-1}\eta$ is a solution of $T\xi=\eta$; if T has a left hand reciprocal T_r^{-1}, then a solution of $T\xi=\eta$, if there exists one, is expressible in the form $\xi=T_l^{-1}\eta$. The following results are well known and find their parallel in the theory of groups.

† See Radon, loc. cit., p. 1384.

(a) If T has both a right hand and a left hand reciprocal, then they are identical and are the only reciprocal of T.

(b) If T has a unique left (or right) hand reciprocal, then this is also a right (or left) hand reciprocal, and it is unique. For if T^{-1} is a left hand reciprocal, then $T^{-1}+STT^{-1}-S$, for every linear limited S on $\mathfrak{S}_2$ to $\mathfrak{S}_1$, is also one, and so $TT^{-1}=I$.

(c) If there is more than one left (or right) hand reciprocal, then there are an infinite number, any linear combination $c_1T_1^{-1}+c_2T_2^{-1}$, with $c_1+c_2=1$, satisfying this condition.

As a consequence we can make the following statements about reciprocals: (1) either there is a unique right hand, or unique left hand reciprocal, and then there is a reciprocal; or (2) there is no right (left) hand reciprocal and an infinite number of left (right) hand reciprocals; or (3) there is neither a left nor a right hand reciprocal.

Obviously the existence of a reciprocal is not essential to the determination of a ξ for a special η such that $T\xi=\eta$.

The convergence of sequences of linear limited transformations may be defined in two ways. (a) For every ξ we have $\lim_n \| T_n(\xi)-T(\xi)\| =0$. By applying the theorem of §2 on sequences of linear operations to $\| T_n(\xi)\|$, it follows that the moduli $M(T_n)$ form a bounded set, and it is obvious that the modulus $M(T)$ is less than or equal to the least upper bound of $M(T_n)$. (b) $\lim T_n=T$ may be defined as equivalent to $\lim_n M(T_n-T)=0$. It is obvious that in this case we have also $\lim_n T_n(\xi)=T(\xi)$ for every ξ, and $\lim_n M(T_n)=M(T)$. If a distinction is necessary one might call the limit in terms of moduli a strong convergence, and the convergence of $T_n(\xi)$ to $T(\xi)$ for every ξ a weak convergence. By using the same type of limit throughout one gets the usual theorems on sums and products of sequences.

Of special interest are the transformations which transform the space $\mathfrak{S}$ into $\mathfrak{S}$ or a subset of $\mathfrak{S}$. In that case, it is possible to consider in addition to T the transformations T^2, T^3, etc., and their linear combinations. In particular we can discuss sequences of transformations defined symbolically by $\sum_{n=1}^{\infty} a_nT^n$, where the a_n's are complex numbers. If the infinite power series $\sum_{n=1}^{\infty} |a_n|\lambda^n$ converges for $\lambda=1/M(T)$, then $\sum_{n=1}^{m} a_nT^n$ converges in the strong sense to the transformation $\sum_{n=1}^{\infty} a_nT^n$ and defines a linear limited transformation. Thus we can obtain

transformations analogous to various analytic functions. For example, $\sum_{n=0}^{\infty} \lambda^n T^n$, $(T^0 = I)$, corresponds to $1/(1-\lambda x)$ and is a linear limited transformation for $|\lambda| < 1/M(T)$; $\sum_{n=0}^{\infty} \lambda^n T^n/(n!)$ corresponds to $e^{\lambda x}$ and gives us a symbolic solution of the linear system of differential equations $d(\xi(\lambda))/d\lambda - T(\xi(\lambda)) = 0$; $\sum_{n=1}^{\infty} \lambda^n T^n/n$ corresponds to $\log(1-\lambda x)$, and so on.

Applying these ideas to the transformation $I-\lambda T$ and denoting by $I-\lambda T_\lambda$ the reciprocal of this transformation for those values of λ for which it exists, we find that this reciprocal exists at least for $|\lambda| < 1/M(T)$ and

$$T_\lambda = -\sum_{n=1}^{\infty} \lambda^{n-1} T^n.$$

Assuming that $T_\lambda = -T$ for $\lambda = 0$, we observe that the relation between T and T_λ is given by

$$T + T_\lambda = \lambda T T_\lambda = \lambda T_\lambda T,$$

the understanding being that the equality holds for all members of $\mathfrak{S}$. Obviously for any two numbers λ and μ for which T_λ and T_μ exist, it is true that

$$T_\mu - T_\lambda = (\lambda - \mu) T_\lambda T_\mu,$$

which includes the preceding relations for $\lambda = 0$ and $\mu = 0$, respectively. It follows at once from this relation that the set of points λ of the complex plane for which the reciprocal T_λ exists form an open set. For if T_{λ_0} exists, then

$$T_\mu = \sum_{n=1}^{\infty} (\lambda_0 - \mu)^{n-1} T_{\lambda_0}^n$$

will satisfy these relations provided $|\mu - \lambda_0| M(T_{\lambda_0}) < 1$. Obviously T_μ is a continuous function of μ in the space of linear limited transformations. These identities give a more general result, namely, if $\lambda_1, \cdots, \lambda_n, \cdots$ is a sequence of values converging to λ_0, such that T_{λ_n} exists for every n, and if $M(T\lambda_n)$ are bounded, then T_{λ_0} exists and $M(T_{\lambda_0} - T_{\lambda_n})$ approaches zero.

On account of the fact that T_λ has the properties of a holomorphic function in the complex λ plane, Riesz† suggests the formation of $\int f(\lambda) T_\lambda$ along continuous curves in the λ plane lying in

† Riesz, *Inf. Inc.*, p. 118.

regions of existence of T_λ, as transformations related to T_λ. He uses with success the transformations $\int T_\lambda \lambda^k d\lambda$ along curves not containing the origin, to separate T into orthogonal sections similar to the decomposition of the kernel of a linear integral equation relative to the singularities of the reciprocal. The method seems worthy of further study.

5. *Completely Continuous Transformations.* The discussion of the singularities of the transformation T_λ and their influence on the transformation T in case T is limited but otherwise unrestricted is somewhat difficult. It is possible, however, to obtain results comparable to those of linear integral equations if the transformation T is restricted to be completely continuous. While a linear limited, and so continuous, transformation transforms a bounded sequence into a bounded sequence and a compact sequence into a compact sequence, a *completely continuous transformation* is defined† to be one which *transforms bounded sequences into compact sequences.* It is obvious that the identity transformation is not completely continuous unless the space $\mathfrak{S}$ is of finite dimension. If we assume for convenience that our transformations are on $\mathfrak{S}$ to $\mathfrak{S}$, then if T and S are two transformations of which one is completely continuous the composites TS and ST are completely continuous, so that the iterates T^2, T^3, $\cdots$ of T are completely continuous if T is. It follows further that, excepting in a space of finite dimension, a completely continuous transformation cannot have a reciprocal. Also if T_n is a sequence of continuous transformations converging to T in the strong sense, that is, $M(T_n - T)$ approaches zero, then T is also completely continuous, which is not true in general if the approach is in the weak sense. The completely continuous transformations form a complete linear subspace of the space of all linear limited transformations, the norm being the modulus.

The simplest completely continuous transformations are those which transform $\mathfrak{S}$ into a subset of finite dimension. It follows that if T is the strong limit of a sequence of such transforma-

† See Riesz, *LFG*, p. 74. Transforming a compact sequence into a bounded sequence makes T still limited, and so continuous. J. von Neumann, (Mathematische Annalen, vol. 102 (1929), p. 70), has suggested a weaker continuity condition in the notion *closed*, T being closed if $\lim \xi_n = \xi$ and $\lim T\xi_n = \eta$ implies $T\xi = \eta$.

tions, then T is completely continuous. Whether the theorem is reversible without additional limitations on the space considered seems to be still undetermined. The converse theorem does hold in the spaces I.1$_e$ ($e \geqq 0$), II.1$_e$ ($e > 0$), and II.3.

Riesz† has shown that the transformations of the form $I-T$ and $I-\lambda T$, where T is completely continuous and λ is a complex number, have properties similar to those of the finite dimensional and Fredholm integral equation case. In particular he has proved the following theorems, mainly on the basis of the fact that for a linear closed set of finite dimension, compactness and boundedness are equivalent.

A. Either the functional equation $\xi - T\xi = \eta$ has a solution ξ for each η of $\mathfrak{S}$ or there exist non-zero solutions of $\xi - T\xi = 0$, the number of linearly independent solutions being finite.

B. There exists a finite integer N such that every solution of $(I-T)^n\xi = 0$ for $n > N$ is a solution of $(I-T)^N\xi = 0$, while $(I-T)^N\xi = 0$ has solutions not present if $n < N$.

C. The transformation $I-T$ transforms $\mathfrak{S}$ into a linear subclass $\mathfrak{S}_1$ of itself. For the same N as in B, the class $\mathfrak{S}_n = \mathfrak{S}_N$ for $n > N$, but $\mathfrak{S}_n > \mathfrak{S}_N$ for $n < N$.

D. There exists a separation of T into two transformations T_1, T_2 which are orthogonal, that is, $T_1T_2 = T_2T_1 = 0$, such that $(I-T_1)(I-T_2) = I-T$; $I-T_2$ has a reciprocal, $I-T_1$ has the same characteristic elements as T.

It is possible to parallel even the adjointness properties as in the Fredholm theory, by noting that the solutions of the homogeneous equation $(I-T)\xi = 0$ can be considered as a transformation T_0 on $\mathfrak{S}$ to the set of solutions T_0 satisfying the conditions

$$(I - T)T_0 = (I - T_0)T_0 = 0.$$

It can be shown that there exists a transformation T_0^* on $\mathfrak{S}$ to a set of finite dimension, the maximum dimension being the same as that belonging to T_0, such that‡

$$T_0^*(I - T_0^*) = T_0^*(I - T) = 0.$$

For the reciprocal $I-\lambda T_\lambda$ of $I-\lambda T$, the singular points are

† *LFG*, pp. 79, etc.

‡ See Hildebrandt, Acta Mathematica, vol. 51 (1928), pp. 311–318.

isolated in the complex plane. It can be shown that the transformation T_λ has an expansion of the form

$$T_\lambda = \sum_{-N}^{\infty} T_n(\lambda - \lambda_0)^n,$$

valid in a neighborhood of the point λ_0, the integer N being identical with the N of properties B and C above, the transformations T_n for $n<0$ being iterates of a transformation S transforming $\mathfrak{S}$ into a linear subspace of finite dimension, thus paralleling the elementary divisor theory.

If now $\lambda_1, \lambda_2, \cdots$ are the singular points of T_λ, we can multiply T_λ by an entire function $D(\lambda)$ having $\lambda=\lambda_n$ as roots of order N_n. The transformation $D(\lambda)\cdot T_\lambda$ will have no singularities in the finite part of the λ plane, and consequently be expressible in the form $\sum_{i=1}^{\infty} T_i\lambda^i$, that is, in essence, T_λ is the quotient of two entire functions, as in the Fredholm theory.

It is possible to extend these results in a number of directions.

(a) Consider a transformation expressed in the form $S-\lambda T$, where S has a reciprocal S^{-1}. Then if the reciprocal of $S-\lambda T$ exists and is written in the form $S^{-1}-\lambda T_\lambda$, the transformation T_λ satisfies reciprocal relations of the form

$$S^{-1}T + T_\lambda S - \lambda T_\lambda T = TS^{-1} + ST_\lambda - \lambda TT_\lambda = 0,$$

which reduce to the usual relations for $S=S^{-1}=I$. If further T is completely continuous, then the properties of $S-\lambda T$ are similar to those of $I-\lambda T$ discussed above.

(b) Returning to the transformation $I-\lambda T$, we see that it possesses a reciprocal at least for $\lambda<1/M(T)$. Let ϵ be the greatest lower bound of values ϵ_1 for which there exists a completely continuous transformation $S_{\epsilon_1}=S$ such that $M(T-S)<\epsilon_1$. If now we write

$$I - \lambda T = I - \lambda(T - S) + \lambda S,$$

then $I-\lambda(T-S)$ has a reciprocal for $|\lambda|<1/\epsilon$. Hence in the circle $|\lambda|<1/\epsilon$ we have a condition similar to the case where T is completely continuous. In particular, there exists a power series $D(\lambda)$ and an expansion $\sum\lambda^n T_n$ such that $D(\lambda)T_\lambda=\sum\lambda^n T_n$, the expansions being valid for $|\lambda|<1/\epsilon$. If $\epsilon=0$ we have the Fredholm or completely continuous case. This is in essence a

generalization of the Schmidt "Abspaltungsverfahren"† the approximating transformation S being completely continuous instead of finite dimensional.‡

6. *Adjoint Spaces and Adjoint Linear Transformations.* In the theory of linear integral equations the adjoint equation, suitably defined, plays an important role. Since on our general range the explicit form of a transformation is not given, it seems a difficult matter to formulate an adjoint theory. Examination of existing instances would indicate that the definition of adjoint transformations might be based on the notion of adjoint spaces, which in turn depends upon the correspondence set up between spaces by linear operations. We consequently formulate, with Hahn,§ the following definition.

The adjoint space $\mathfrak{T}$ of any linear complete vector space $\mathfrak{S}$ consists of the totality of all linear limited operations on $\mathfrak{S}$.

It is obvious that the space $\mathfrak{T}$ is linear. If we take as norm the modulus of the operation, then the norm satisfies the triangle property, and the space is complete. If we denote elements of $\mathfrak{T}$ by η, then every linear limited operation on $\mathfrak{S}$ is expressible in the form

$$L(\xi) = (\eta, \xi)$$

where (η, ξ) is a bilinear operation on $\mathfrak{T}\mathfrak{S}$ whose modulus is unity. Obviously, for fixed ξ, (η, ξ) is a linear operation on $\mathfrak{T}$.

The adjoint space of any given space may be realized in more than one way, but the different realizations are equivalent in the sense that there exists a one to one correspondence between them, that is, if $\mathfrak{T}_1$ and $\mathfrak{T}_2$ are adjoint to $\mathfrak{S}$, then there exists a linear limited transformation on $\mathfrak{T}_1$ to $\mathfrak{T}_2$ with a reciprocal.

As examples of adjointness, we note the following. The Hil-

† For a similar procedure, see Radon, Wiener Berichte, vol. 128 IIa (1919), pp. 1106–1114; Hille and Tamarkin, Annals of Mathematics, vol. 31 (1930), p. 509.

‡ We might mention that, while the first general theory of integral equations of E. H. Moore (this Bulletin, vol. 28 (1912), pp. 334–362) does not seem to be a special case under the above developments, still he treats completely continuous transformations, particularly of the type which can be uniformly approximated by transformations of finite dimension. The non-determinantal results can in fact be derived, following Schmidt, with fewer postulates than are sufficient to make Fredholm determinants available.

§ Journal für Mathematik, vol. 157 (1927), p. 218.

bert space $I.1_1$ and $\mathfrak{L}^2$ space $II.1_1$ are self-adjoint; the adjoints of $I.1_e$ and $II.1_e$ $(e>0)$ are $I.1_{1/e}$ and $II.1_{1/e}$, respectively; the adjoint of $I.1_0$ is I.4, of $I.3_0$ is $I.1_0$, and so on.

The adjoint of the adjoint space $\mathfrak{T}$ of $\mathfrak{S}$ while including $\mathfrak{S}$ may be larger. Thus the adjoint of $I.3_0$ is $I.1_0$, of $I.1_0$ is I.4 and so on. How often this process of taking adjoints may be repeated is still undetermined. Hahn calls a space where the adjoint of $\mathfrak{T}$ is $\mathfrak{S}$ *regular*. $I.1_e$ and $II.1_e$ $(e>0)$ are regular.

The notion of adjoint space makes possible the definition of a weak limit in $\mathfrak{S}$. Thus weak limit $\xi_n=\xi$ may be defined as

$$\lim (\eta, \xi_n) = (\eta, \xi) \text{ for every } \eta \text{ of } \mathfrak{T}.$$

Since the adjoint of $\mathfrak{T}$ may be other than $\mathfrak{S}$, the definition of a weak limit in $\mathfrak{T}$, namely,

$$\text{weak limit } \eta_n=\eta, \text{ equivalent to } \lim (\eta_n, \xi)=(\eta, \xi) \text{ for every } \xi,$$

would in general be still weaker. One could write down a weak limit relative to a complete subclass $\mathfrak{T}_0$ of $\mathfrak{T}$, the η ranging over this class only. It is obvious that if weak lim $\xi_n=\xi$, then $\|\xi_n\|$ form a bounded sequence. If the space $\mathfrak{S}$ is the adjoint of $\mathfrak{T}_0$ and is separable, then by applying the theorem† on the compactness of sets on the denumerable range, it is possible to show that every bounded subset of $\mathfrak{S}$ is weakly compact.

With the aid of the bilinear operation (η, ξ) it is possible to define the notion of adjoint transformation. If T is a linear limited transformation on $\mathfrak{S}_1$ to $\mathfrak{S}_2$ and η_2 any element of $\mathfrak{S}_2$,* the adjoint of $\mathfrak{S}_2$, then $(\eta_2, T(\xi_1))$ defines a linear limited operation of $\mathfrak{S}_1$, and is consequently expressible in the form (η_1, ξ_1), where η_1 belongs to $\mathfrak{T}_1$. We have thus defined a transformation T^* on $\mathfrak{T}_2$ to $\mathfrak{T}_1$, which is obviously linear and limited and has the same modulus as T. The adjoint of T^* is T if $\mathfrak{S}_1$ is regular, otherwise it agrees with T on $\mathfrak{S}_1$, but may be on a space including $\mathfrak{S}_1$ to a space including $\mathfrak{S}_2$.

If T_1 and T_2 have as adjoints T_1^* and T_2^*, then the adjoint of T_1T_2 is $T_2^*T_1^*$. Further the adjoint of the identity is the identity. Hence if T has a reciprocal, the adjoint T^* also has a reciprocal, namely, $(T^{-1})^*$.

If T transforms the space $\mathfrak{S}$ into a space of finite dimensions,

† See Fréchet, Rendiconti di Palermo, vol. 26 (1906), p. 42.

the adjoint T^* on $\mathfrak{T}$ has the same property, the dimensions of the space being the same for both.

If the space $\mathfrak{S}$ is separable, then the adjoint of a completely continuous transformation on $\mathfrak{S}$ to $\mathfrak{S}$ is again completely continuous. Whether this holds in general, is undetermined. However if T is completely continuous, then the transformation $I-\lambda T^*$ has properties similar to those of $I-\lambda T$. In particular, the transformation $I-\lambda T^*$ has a reciprocal if and only $I-\lambda T$ has, the equation $I-\lambda T^*=0$ has as many linearly independent solutions as $I-\lambda T=0$, and the reciprocal $I-\lambda T_\lambda{}^*$ of $I-\lambda T^*$ is expressible in the form

$$\left(\sum_{i=0}^{\infty} d_i\lambda^i\right) T_\lambda^* = \sum_{i=0}^{\infty} T_i^*\lambda^i,$$

where $\sum d_i\lambda^i$ is a permanently convergent power series, and $\sum\lambda^i T_i^*$ converges for all values of λ in the modular sense, the T_i^* being the adjoints of the transformations corresponding to T_λ.

It is a simple matter to prove the following results due to Riesz:

(a) in order that for a given ξ_2 of $\mathfrak{S}_2$ there exist a ξ_1 of $\mathfrak{S}_1$ such that

$$T(\xi_1) = \xi_2,$$

it is necessary that there exist a constant M such that for every η_2 of $\mathfrak{T}_2$

$$|(\eta_2, \xi_2)| \leqq M\|T^*(\eta_2)\|;$$

and (b) in order that T have a reciprocal, it is necessary that there exist an M such that for every η_1 of $\mathfrak{T}_1$ and ξ_1 of $\mathfrak{S}_1$

$$|(\eta_1, \xi_1)| \leqq M\|T(\xi_1)\| \cdot \|\eta_1\|,$$

and for every η_2 of $\mathfrak{T}_2$ and ξ_2 of $\mathfrak{S}_2$

$$|(\eta_2, \xi_2)| \leqq M\|T^*(\eta_2)\| \cdot \|\xi_2\|.$$

Riesz and Radon† have shown that these conditions are also sufficient in special cases, characterized, perhaps, by the fact

† See Riesz, Mathematische Annalen, vol. 69 (1910), pp. 469 etc.; Annales de l'École Normale, (3), vol. 28 (1911), pp. 33–62; Riesz, *Inf. Inc.*, Chap. IV; Radon, Wiener Berichte, vol. 122, IIa (1913), pp. 1386, etc.; Helly, Monatshefte, vol. 31 (1921), pp. 60–91.

that the spaces $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are separable, and have regular adjoint spaces.

7. *Hilbert or Hermitian Spaces. Hermitian Transformations.*† The derivation of more detailed results on the singularities of the transformations T and $I-\lambda T$ is hampered by a lack of symmetry in the general vector space which forms the basis of the preceding sections. Limiting the spaces to which our considerations apply by introducing an element of symmetry into the bilinear operation defining adjoint spaces makes it possible not only to obtain a larger body of results for limited transformations, but even to consider special types of transformations, unlimited, or defined only on a portion of the space. The principal additional conditions as given by J. V. Neumann‡ and M. H. Stone§ on the space are in connection with condition B on our vector space. The norm $\|\xi\|$ is defined in terms of a bipartite operation (ξ, η) on the space to complex numbers, subject to the Hermitian condition

$$(\xi, \eta) = (\overline{\eta, \xi}),$$

linear in the first element and so conjugately linear in the second, and such that for every ξ, $(\xi, \xi) \geqq 0$, zero only if $\xi = 0$. Then by definition

$$\|\xi\|^2 = (\xi, \xi),$$

and this satisfies the conditions on norm. In addition it is assumed (D) that the space is separable, that is, a denumerable set is dense in the space. A complete vector space satisfying these conditions if of infinite dimensions is called a Hilbert space. Due to the fact that after all the Hermitian operator (ξ, η) is central, it seems more to the point to call it a Hermitian

† The principal references for this section are the following: Hellinger-Toeplitz, Encyklopädie der Mathematischen Wissenschaften, vol. IIC, p. 1575; Riesz, *Inf. Inc.*, Chapters IV and V; E. H. Moore, Mathematische Annalen, vol. 86 (1922), pp. 34; T. Carleman, *Sur les équations intégrales singulières*, Upsala Arsskrift, 1923; J. v. Neumann, Mathematische Annalen, vol. 102 (1929), pp. 49–137, 370–427; M. H. Stone, Proceedings of the National Academy, vol. 15 (1929), pp. 198–200, 423–5; A. Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929; F. Riesz, Szeged Acta Litterarum, vol. 5 (1930), pp. 23–54; this will be cited in what follows as Szeged Acta.

‡ Loc. cit., p. 64.

§ Loc. cit., p. 198.

space, reserving the name Hilbert space for the space $I.1_1$, that is, $\sum |x_n|^2 < \infty$. In any case we shall denote by $\mathfrak{H}$ a space which is (A) a vector space, (B) such that $\|\xi\| = (\xi, \xi)$, (C) complete, and (D) separable.

The usual instances of such spaces are the Hilbert space $I.1_1$ with $(\xi, \eta) = \sum x_p \bar{y}_p$ and the $\mathfrak{L}^2$ space $II.1_1$ with $(\xi, \eta) = \int \xi \bar{\eta}$. E. H. Moore† has given a means of constructing such spaces, not all of which satisfy the separability condition D, however. If $\mathfrak{P}$ is any class of elements and $\epsilon(p, q)$ a complex-valued function satisfying the condition that for any finite set $p_1, \cdots, p_n$ of elements of $\mathfrak{P}$, the matrix $\epsilon(p_i, p_j)$ is the matrix of coefficients of a positive Hermitian form, then the function $\xi(p)$ is a member of the space $\mathfrak{S}$ if there exists a constant N such that for every finite set of elements $p_1, \cdots, p_n$ and for any set of complex numbers $x_1, \cdots, x_n$ we have

$$\sum_{i=1}^{n} \bar{x}_i \xi(p_i) \sum_{j=1}^{n} x_j \bar{\xi}(p_j) \leqq N \sum_{i,j=1}^{n} \bar{x}_i \epsilon(p_i p_j) x_j.$$

The greatest lower bound of the possible numbers N is $\|\xi\|^2$. The operation (ξ, η) is

$$\lim_\sigma \det \begin{pmatrix} \epsilon(p_i p_j) & \xi(p_i) \\ \eta(p_j) & 0 \end{pmatrix} \div \det \epsilon(p_i p_j)$$

where σ is any set $p_1, \cdots, p_n$ of elements of $\mathfrak{P}$ for which $\det \epsilon(p_i p_j)$ is not zero and $\lim_\sigma F(\sigma) = a$ is defined; for every $e > 0$ there exists a σ_e such that for any σ containing all the points of σ_e we have $|F(\sigma) - a| \leqq e$. Interesting instances of these spaces are (1) $\mathfrak{P} = [0 \leqq p \leqq 1]$; $\epsilon(p, q) = \delta(pq) = 0$ for $p \neq q$ and 1 for $p = q$; (2) in the space of infinite dimensions, consider a matrix set of elements a_{pq} such that $\sum_q |a_{pq}|^2 < \infty$; then $\epsilon(p, q) = \sum_r \bar{a}_{pr} a_{qr}$ gives a space of interest in connection with the Schmidt‡ solution of the system

$$\sum_q a_{pq} x_q = y_p$$

for which $\sum |x_p|^2 < \infty$; (3) the analogue of (2) in $\mathfrak{L}^2$ space, namely, $f_n(p)$ such that $\int |f_n|^2 < \infty$ and $\epsilon(n, m) = \int \bar{f}_n f_m$§; (4) if

† Loc. cit., pp. 34, etc.

‡ See Rendiconti di Palermo, vol. 25 (1908), pp. 53–77; Riesz, *Inf. Inc.*, pp. 65, etc.

§ See A. J. Pell, loc. cit.

$\mathfrak{P}$ is an additive class of sets, and η is a positive absolutely additive set function on $\mathfrak{P}$, then $\epsilon(p, q) = \epsilon(E_1, E_2) = \eta(E_1 \cdot E_2)$ will give rise to integrals of the Hellinger-Radon type.

Returning to the space $\mathfrak{H}$, in terms of (ξ, η), we have a Schwarz inequality

$$(\xi, \eta)(\eta, \xi) \leqq (\xi, \xi)(\eta, \eta) = \|\xi\|^2 \|\eta\|^2.$$

Further it is possible to define orthogonality, to reduce any system of elements to a linearly equivalent system of normed orthogonal elements, necessarily denumerable, and to express any element ξ in terms of a complete system of normed orthogonal elements (that is, one whose linear extension is dense in $\mathfrak{H}$), by the Fourier expansion

$$\xi = \sum_{n=1}^{\infty} (\xi, \phi_n)\phi_n,$$

(convergence in the sense of norms) with

$$(\xi, \xi) = \sum_{n=1}^{\infty} (\xi\phi_n)(\phi_n\xi) = \sum_{n=1}^{\infty} |(\xi, \phi_n)|^2.$$

As a consequence it is possible to set up a one to one correspondence between any space $\mathfrak{H}$ and the Hilbert space $I.1_1$.†

Riesz‡ has pointed out that the following lemma is still valid even if the separability condition D is not fulfilled:

LEMMA I. *If $\mathfrak{H}_0$ is any subset of $\mathfrak{H}$ which is not dense in $\mathfrak{H}$, then there exists an element ξ of $\mathfrak{H}$, for which $\|\xi\| = 1$, which is orthogonal to all the elements of $\mathfrak{H}_0$.*

By the use of the Fourier expansion it is possible to make the connection between the adjoint operation of the preceding section and (ξ, η) as follows:

LEMMA II.§ *If $L(\xi)$ is a linear limited operation on $\mathfrak{H}$, or on*

† If the separability condition D were dropped, a non-denumerable set of mutually orthogonal functions would be possible, for example, in the Moore instance (1). However any given element would be orthogonal to all but a denumerable subset of these, and the Fourier expansion would still be valid in form. Compare also the situation in almost periodic functions.

‡ Szeged Acta, p. 28.

§ See also J. von Neumann, loc. cit., p. 94.

a subset of $\mathfrak{H}$ dense in $\mathfrak{H}$, then there exists a unique element η of $\mathfrak{H}$ such that

$$L(\xi) = (\xi, \eta).$$

8. *Linear Limited Transformations.* When it comes to linear limited transformations T on $\mathfrak{H}$ to $\mathfrak{H}$, it is obvious that any complete normed orthogonal system will set up a one to one correspondence between T and a matrix t_{pq}, formed from the coefficients of a linear limited transformation on Hilbert space I.1$_1$ to the same space.† As a consequence all of the results obtained by Hilbert and his followers for this theory can be translated to the space $\mathfrak{H}$. However, some of the results gain in elegance and lucidity by being derived without using the explicit form of the transformation on the space $\mathfrak{H}$, and point the way to further generalizations.

Since every linear limited operation on $\mathfrak{H}$ is of the form (ξ, η) we can at once define the adjoint of a linear limited transformation T by the existent T^* such that for every ξ and η

$$(T\xi, \eta) = (\xi, T^*\eta).$$

If $T = T^*$, then T is said to be self adjoint or Hermitian, and we shall denote such a transformation by H. Obviously, for any ξ, $(H\xi, \xi)$ is a real number. If for all ξ we have $(H\xi, \xi) \geqq 0$, H is said to be positive, negative being similarly defined. Connected with any limited transformation T are the Hermitian transformations

$$H_1 = T + T^*, \quad H_2 = i(T - T^*), \quad H_3 = T^*T \quad \text{and} \quad H_4 = TT^*.$$

In terms of H_1 and H_2 we obviously have

$$T = \tfrac{1}{2}(H_1 - iH_2).$$

The transformations H_3 and H_4 are obviously positive. In case H_3 and H_4 are the same, that is, $TT^* = T^*T$, the transformation T is said to be normal, for which a necessary and sufficient condition is that H_1 and H_2 be commutative. If finally $TT^* = T^*T = I$, then T is the analogue of an orthogonal transformation,

† A similar correspondence is obviously possible if T is on a space $\mathfrak{H}$ to another space $\mathfrak{H}'$, in which $\mathfrak{H}$ and $\mathfrak{H}'$ might agree except for the fundamental normed orthogonal system used.

and is called unitary (U). Obviously $\|U\xi\| = \|\xi\|$, that is, U leaves lengths of vectors unchanged.

If H is limited, then the set of values $(H\xi, \xi)$ for $\|\xi\| = 1$ is limited to a portion of the real axis: (m, M). For the case of any linear transformation T, not necessarily limited, but defined on the entire space $\mathfrak{H}$, Stone† has generalized a theorem of Hausdorff on forms in n variables, to the effect that the set of complex values of $(T\xi, \xi)$ subject to the condition $\|\xi\| = 1$ forms a convex set of points in the complex plane.

Toeplitz has shown that if T is a transformation of the form $\sum_q t_{pq}x_q$ on $I.1_1$ space to $I.1_1$ space, and valid for all points of the space, then it is necessarily limited.‡ In dealing with a transformation on the space $\mathfrak{H}$, assumed to be linear only, it does not follow that it can be realized by a transformation of the form indicated in the space $I.1_1$, that is, the Toeplitz theorem is not immediately transferable. In this direction there is a result of J. von Neumann,§ namely, that if T is a Hermitian transformation on all of $\mathfrak{H}$ space and satisfies the weaker closed condition that if ξ_n approaches ξ and $T\xi_n$ approaches η, then $T(\xi) = \eta$, it follows that T is limited.

Because of their greater simplicity, the properties of linear limited Hermitian transformations are most elegant and extensive. The following theorem is of particular importance.

THEOREM 2. *The transformation $I - \lambda H$ has a reciprocal for all complex and real λ exterior to the interval (m, M) of variation of $(H\xi, \xi)$ for $\|\xi\| = 1$.*

Let $\lambda = \lambda_1 + i\lambda_2$. Then if λ is exterior to (m, M),

$$\|(I - \lambda H)\xi\|^2 = \|\xi\|^2 - 2\lambda_1(H\xi, \xi) + (\lambda_1^2 + \lambda_2^2)\|H\xi\|^2 > K\|\xi\|^2,$$

where K is a positive number depending on λ. It follows from this inequality that $(I - \lambda H)\xi = 0$ has as unique solution $\xi = 0$, and that $(I - \lambda H)$ operating on the space $\mathfrak{H}$ gives a closed subset $\mathfrak{H}_0$. The fact that this set is dense on $\mathfrak{H}$ and so agrees with $\mathfrak{H}$ now follows at once from Lemma 1. For if η were such that $\|\eta\| = 1$ and were orthogonal to $\mathfrak{H}_0$, then $(H\eta, \eta)$ would take on a value λ exterior to (m, M).

† This Bulletin, vol. 36 (1930), pp. 259–62.

‡ Riesz, *Inf. Inc.*, pp. 78, 82, suggest that this is true for other spaces of Group I and is probably deducible from the general theorem of §3 above.

§ Loc. cit., p. 107.

As a consequence of this theorem and the results on reciprocals of §4, it follows that if H is such that the values of $(H\xi, \xi)$ for $\|\xi\| = 1$ are bounded from zero, then there exists a number $k \neq 0$ such that $kH = I - (I - kH)$ and so H has a reciprocal.† An immediate consequence of this is that if, for any limited transformation T, $(TT^*\xi, \xi)$ and $(T^*T\xi, \xi)$ are bounded from zero for $\|\xi\| = 1$, then TT^* and T^*T have reciprocals, and T has a reciprocal, namely,

$$(T^*T)^{-1}T^* = T^*(TT^*)^{-1}.$$

These conditions are also necessary.‡

The next result is fundamental.§

THEOREM 3. *To every linear limited transformation H there correspond two transformations E_- and E_+ (Einzeltransformationen) having the following properties*:

(1) $\quad E_-E_- = E_-, \quad E_+E_+ = E_+, \quad I = E_- + E_+.$

(2) *E_- and E_+ are commutative and commutative with any transformation commutative with H.*

(3) *HE_- and HE_+ are respectively negative and positive.*

(4) *For all ξ such that $H\xi = 0$ we have $E_-\xi = 0$ and $E_+\xi = \xi$.*

At first sight it might seem that $E_+\xi$ could be defined as ξ if $(H\xi, \xi) \geqq 0$ and zero if $(H\xi, \xi) \leqq 0$, and $E_-\xi = \xi - E_+\xi$. However, the transformations so defined are not necessarily linear, nor do they satisfy condition (2). Instead the path leads through the notion of polynomials in H, and via uniformly convergent sequences of polynomials to continuous functions in H corresponding to polynomials and continuous functions on (m, M), positive if the corresponding function is positive and mutually commutative. The transformation $H_+ = HE_+$ corresponds to $\frac{1}{2}(t + |t|)$ and $H_- = HE_-$ to $\frac{1}{2}(t - |t|)$. Then H_+ and H_- are orthogonal. The transformation E_- is then defined as transforming all of the functions of the set $[H_-\xi]$ into themselves, and the functions orthogonal to these into zero, while we set $E_+ = I - E_-$. If the parallelism between functions and transfor-

† Method due to Hilb; see Riesz, *Inf. Inc.*, pp. 93, etc.

‡ See, for instance, Wintner, loc. cit. p. 138.

§ See J. von Neumann, loc. cit., pp. 112–4; Riesz, Szeged Acta, pp. 31–38; *Inf. Inc.*, pp. 135–138.

mations is extended to certain discontinuous functions, then E_- corresponds to the function unity for $t<0$ and zero for $t\geqq 0$.

Let now $E(\lambda)$ be the transformation of the type E_- corresponding to the Hermitian transformation $I-\lambda H$, λ real. These transformations correspond to a fundamental set of characteristic functions. They have the following properties:

(1) $E(\lambda)\xi=0$ if $\lambda<m$; $E(\lambda)\xi=I\xi=\xi$ if $\lambda>M$.

(2) $E(\lambda_1)E(\lambda_2)\xi=E(\lambda_1)\xi$ if $\lambda_1\leqq\lambda_2$ so that $(E(\lambda)\xi, \xi)$ is increasing in λ for each ξ.

(3) $E(\lambda)\xi$ is continuous on the left.

$$(4)\quad H\xi=\int_m^M \lambda d_\lambda E(\lambda)\xi, \qquad I\xi=\int_m^M d_\lambda E(\lambda)\xi,$$

the integral being taken in the sense of Stieltjes extended to vector space and norm convergence. The set $E(\lambda)$ is uniquely determined by the transformation H. It is called by Stone a *canonical resolution of the identity*. Riesz suggests the term *spectral set*.

The transformations $E(\lambda)$ define the spectrum of $I-\lambda H$, and make available a classification of the points of the interval (m, M). For instance the points λ_0 for which there exists a ξ with $\|\xi\|=1$ and for which $(E(\lambda)\xi, \xi)$ is discontinuous at λ_0 correspond to the point spectrum, which, because of the separability of $\mathfrak{H}$, forms a denumerable set. For such λ_0 the homogeneous equation $\xi-\lambda_0 H\xi=0$ has a non-zero solution. By subtracting, so to speak, the discontinuities of $E(\lambda)$ one gets the part which gives the continuous spectrum and the differential solutions of the homogeneous equation.†

If $f(H)$ corresponds to the continuous function $f(x)$ on (m, M), then

$$f(H)\xi=\int_m^M f(\lambda)d_\lambda E(\lambda)\xi.$$

In particular, the notion is extensible to non-Hermitian transformations associated with H, and we have for the reciprocal

$$(I-\mu H)^{-1}=\int_m^M \frac{d_\lambda E(\lambda)\xi}{\lambda-\mu},$$

† For details, see Riesz, *Inf. Inc.*, pp. 141, etc.

for the set of values of μ for which it exists.

Similar results apply to the normal limited transformations T, for which $T^*T=TT^*$, and which are expressible in the form H_1+iH_2, H_1 and H_2 being commutative and Hermitian. If $E_1(\lambda)$ corresponds to H_1 and $E_2(\mu)$ to H_2, then E_1 and E_2 are commutative for every λ and μ. By utilizing the fact that $I=\int dE(\lambda)$, we have

$$T = H_1 + iH_2 = \int_{m_2}^{M_2}\int_{m_1}^{M_1} (\lambda + i\mu) d_{\lambda\mu} E_1(\lambda) E_2(\mu)$$

$$= \int_{m_1}^{M_1}\int_{m_2}^{M_2} (\lambda + i\mu) d_{\mu\lambda} E_2(\mu) E_1(\lambda),$$

where $\Delta_{\lambda\mu}E_1(\lambda)E_2(\mu) = (E_1(\lambda+\Delta\lambda)-E_1(\lambda))(E_2(\mu+\Delta\mu)-E_2(\mu))$. In a sense the normal transformation is the plane analogue of the Hermitian.

In particular if T is unitary U, then $I-\lambda U$ has a reciprocal for all values of λ not on the unit circle in the complex plane. Utilizing this fact, and performing a transformation to polar coordinates, we see that to every unitary transformation there corresponds a spectral set $E(\lambda)$ on the interval $(0, 2\pi)$ such that

$$U\xi = \int_0^{2\pi} e^{i\lambda} d_\lambda E(\lambda)\xi$$

and conversely. This suggests that the unitary transformations are connected with the Hermitian by a transformation of the symbolic form $U=e^{iH}$.†

Obviously this method of representation can be extended to multiplicative combinations of Hermitian, unitary, and normal transformations.

9. *Unlimited Transformations.* After Hilbert the first steps in the discussion of such transformations were taken by Carleman. The introduction of infinite matrices as a tool in atomic theory gave a new impulse towards the investigation of such transformations. So far the chief contributors have been von Neumann, Stone, and Wintner.

The transformations so far treated have been limited and de-

† See J. von Neumann, loc. cit., pp. 119, etc.; Wintner, Mathematische Zeitschrift, pp. 268, etc.

fined for the entire space. If we drop the second of these conditions, then it seems natural to demand at least that a transformation be defined and linear on a linear subset $\mathfrak{F}$ dense in $\mathfrak{H}$. (For example, id/dx on the class of polynomials, dense in continuous function set.) If in addition the transformation T is limited on a set $\mathfrak{F}$ dense in $\mathfrak{H}$, it is uniformly continuous on this set and so can be extended to a continuous, that is, a limited transformation on $\mathfrak{H}$. Consequently if a transformation is not limited, there exists a subset of $\mathfrak{H}$ dense in $\mathfrak{H}$, on which it is not limited.

To define the adjoint of a transformation T, we select the elements η of $\mathfrak{H}$ for which $(T\xi, \eta)$ is a linear limited operation on $\mathfrak{F}$, so that there exists an element η^* such that $(T\xi, \eta) = (\xi, \eta^*)$ for all ξ of $\mathfrak{F}$. We thus define $T^*\eta = \eta^*$ on $\mathfrak{F}^*$. If $\mathfrak{F}^*$ is dense on $\mathfrak{H}$, then we can similarly define T^{**} and $\mathfrak{F}^{**}$, and $\mathfrak{F}^{**}$ will include $\mathfrak{F}$. If in particular $\mathfrak{F}^* = \mathfrak{F}$ and $T^* = T$, then T is said to be Hermitian (von Neumann: hypermaximal; Stone: self-adjoint) and denoted by H. A Hermitian transformation is said to be commutative with a limited transformation T if both ξ and $T\xi$ belong to $\mathfrak{F}$ and $HT\xi = TH\xi$ on $\mathfrak{F}$.

As in the case of limited transformations, if H is Hermitian on $\mathfrak{F}$, then for all λ which are not real the transformation $I - \lambda H$ possesses a reciprocal which is a limited transformation on $\mathfrak{H}$. In particular $I + iH$ and $I - iH$ have reciprocals B and B^*, which are adjoint. Riesz† observes that if one defines E_- and E_+ for the transformation

$$D = B + B^* = 2B^*HB\,,$$

then these will serve also as the E_- and E_+ for the transformation H, so that it is possible to define the transformations $E(\lambda)$ for the transformation $I - \lambda H$, λ being real. The final result is as follows.

THEOREM 4. *Every Hermitian transformation gives rise to a spectral set $E(\lambda)$, defined on $(-\infty, \infty)$, zero for $\lambda = -\infty$ and the identity for $\lambda = \infty$, such that a necessary and sufficient condition that ξ belong to $\mathfrak{F}$, that is, H be defined for ξ, is that the Stieltjes integral*

$$\int_{-\infty}^{\infty} |\lambda|^2 d_\lambda (E(\lambda)_-, \xi)$$

† Szeged Acta, p. 44.

exist, the integral being defined as

$$\lim_{\max\Delta_i\lambda\to 0} \sum_{i=-\infty}^{\infty} |l_i|^2 \Delta_i(E(\lambda)\xi, \xi).$$

Then $H\xi = \int_{-\infty}^{\infty} \lambda d_\lambda E(\lambda)\xi$, *existent in the sense of norms, and*

$$(I - \mu H)^{-1} = \int_{-\infty}^{\infty} \frac{d_\lambda E(\lambda)}{\lambda - \mu}$$

where it exists.

J. von Neumann† obtains the same result another way. Essentially, his method involves the observation that $I+iH$ and $I-iH$ having limited reciprocals, each transforms $\mathfrak{F}$ into $\mathfrak{H}$, and therefore they define a limited transformation U which is unitary. Solving the symbolic equation

$$U(I + iH) = I - iH,$$

for H and applying the result to the canonical representation of a U, namely, $\int_0^{2\pi} e^{i\lambda} dE(\lambda)$ or $\int_{-\pi}^{\pi} e^{i\lambda} dE(\lambda)$, suggests the transformation $\lambda = 2 \arctan \rho$, which leads to the same final characterization of the field $\mathfrak{F}$ and the transformation H. Stone‡ obtains the same result by working through a representation for the reciprocal of $I-\lambda H$ for λ not real.

It is perhaps interesting to note that the sums involved in defining the set $\mathfrak{F}$ of elements ξ for which H is defined are analogous to the one involved in defining the set of Lebesgue integrable square functions, the value $(E(\lambda)\xi, \xi)$ replacing the measure of the set E for which $f \leqq \lambda$; further that the set $\mathfrak{F}$ of elements ξ form again a Hilbert space, the Hermitian operator being

$$[\xi, \eta] = \int_{-\infty}^{\infty} |\lambda|^2 d_\lambda(E(\lambda)\xi, E(\lambda)\eta),$$

but the elements $H\xi$ do not necessarily belong to this space.

J. von Neumann has investigated further transformations which satisfy the Hermitian condition for only a part of the set on which they are definable. For these the set $\mathfrak{F}^*$ would con-

† Loc. cit., pp. 80, 92.

‡ Loc. cit., pp. 423–425.

tain $\mathfrak{F}$, and on $\mathfrak{F}$ we have $T\xi = T^*\xi$, but the transformation T^* is not Hermitian, that is, there exists an η of $\mathfrak{H}$ for which $(\eta, T^*\eta)$ is not real. If $\mathfrak{F}$ is the set on which $(\eta, T^*\eta)$ is real, then T is said to be maximal. He finds by utilizing the correspondence defined by $I+iT$ and $I-iT$, called the Cayley transform, that the maximal transformations are expressible as the sum of a Hermitian transformation and a finite or denumerable set of transformations for which the Cayley transformation has for a certain normed orthogonal system of elements $\phi_1, \cdots, \phi_n, \cdots$ the form

$$U\phi_n = \phi_{n-1}.$$

J. von Neumann has considered further the transformations which as given are neither Hermitian nor maximal, and finds very interesting properties, particularly with respect to a certain amount of indefiniteness in the singularities of their reciprocals, first observed by Carleman.

It is to be hoped that the impetus given by these recent researches will be effective in encouraging further investigation into the characteristics of general limited and unlimited transformations in $\mathfrak{H}$ spaces and general vector spaces.

THE UNIVERSITY OF MICHIGAN

ON THE ZEROS OF EXPONENTIAL SUMS AND INTEGRALS*

BY R. E. LANGER

1. *Introduction.* It is of frequent occurrence in problems of both pure and applied mathematics that certain values sought may be specified and must be determined as the roots of a transcendental equation. In particular, the equation may be of the class in which the unknown is involved only through the medium of exponential or trigonometric functions, with coefficients which are power functions or essentially such. As simple examples one has such equations as

$$\tan z - z = 0, \quad e^z + z^2 e^{3z} + (1 - z) = 0,$$

$$\sin z + z^3 \tan 5z = 0, \quad e^{iz} - 2z^5 = 0.$$

A determination of the distribution of roots of such an equation may, therefore, be of interest and importance. Inasmuch as this problem and certain of its generalizations have been the subjects of recent and successful investigations I have presumed that an outline of methods and attainments associated with it might not be unsuitable for presentation at this meeting of the Society.

The relation between the trigonometric functions of z and the function e^{iz} evidently makes possible the expression of each of the functions involved in the equations above in the form

$$\Phi(z) = \sum_{j=0}^{n} A_j(z) e^{c_j z}. \tag{1}$$

This type of function we shall refer to as an exponential sum. It is to be the subject of Part I of the discussion to follow. If the values $A_j(z)$ and c_j are constants, and the c_j are real, the function (1) can be expressed with the use of an appropriate constant z_0 and a suitable step-function $\phi(t)$ in the form

* An address delivered at the meeting of the Society at Columbia, Missouri, November 29, 1930, by invitation of the program committee.

$$\Phi(z - z_0) = z \int_{c_0}^{c_n} \phi(t) e^{tz} dt. \tag{2}$$

The integral of the type here involved with less specialized function $\phi(t)$ represents, therefore, a generalization of certain sums of type (1). Such integrals include a number of standard functions of importance, among which the Bessel functions may be specifically mentioned. They will be considered briefly in part II of this discussion.

Returning to the consideration of the sum (1) we shall understand that the quantities c_j, to be called the exponents, are constants (real or complex). The functions $A_j(z)$, to be called the coefficients, will be characterized as the discussion progresses by various sets of hypotheses. It will be my general plan to pass in the considerations from forms of less to forms of greater generality, and to indicate at each stage the results and their relation with those previously established.

Part I. The Exponential Sum

2. *Constant Coefficients and Real Commensurable Exponents.* When the exponents c_j in a sum of the form (1) are real, it is a matter merely of the arrangement of terms to suppose that these exponents occur in the order of increasing algebraic magnitude. With this arrangement it is then permissible without loss of generality to assume that the first exponent c_0 is zero, inasmuch as the removal of any exponential factor can in no way affect either the number or distribution of the zeros of the function. These adjustments will be assumed to have been made. Theoretically the simplest type of sum is one, then, in which the problem of the distribution of zeros is essentially an algebraic one, and this occurs in particular when the coefficients are constants ($A_j(z) \equiv a_j$), and the exponents are commensurable, that is, $c_j = \alpha p_j$, $j = 1, 2, \cdots, n$, with some real constant α and integral values p_j. The sum is then of the form

$$\Phi(z) = \sum_{j=0}^{n} a_j (e^{\alpha z})^{p_j}, \quad p_0 = 0, \tag{3}$$

which is a polynomial of degree p_n in the quantity $e^{\alpha z}$. If this polynomial admits as zeros the values $\xi_1, \xi_2, \cdots, \xi_{p_n}$, the func-

tion (3) vanishes for such values of z, and for only such, as satisfy a relation

$$e^{\alpha z} = \xi_j.$$

The zeros of $\Phi(z)$ are, therefore, given by the formula

$$z = \frac{1}{\alpha}\{2m\pi i + \log \xi_j\}, \qquad \begin{matrix}(j = 1, 2, \cdots, p_n),\\ (m = 0, \pm 1, \pm 2, \cdots).\end{matrix} \tag{4}$$

They are thus seen to be denumerably infinite in number and to be distributed in the complex z plane at regular intervals of length $2\pi/\alpha$ along p_n lines (not necessarily distinct) which are normal to the axis of reals.

If the explicit solution of the polynomial equation involved is feasible, the determination of the zeros of $\Phi(z)$ is completed by the formula (4). In the alternative case the following further considerations will not be futile. It is clear from the result (4), that with any specifically given function $\Phi(z)$, the choice of a constant K is possible so that the zeros of $\Phi(z)$ all lie within the rectilinear strip of the z plane given by the relation

$$|x| < K, \qquad (z = x + iy). \tag{5}$$

Any line $y = y_1$, parallel to the axis of reals, cuts this strip, and on such a line it is found by direct computation that

$$I\left\{\frac{1}{a_0}\Phi(z)\right\} = \sum_{j=1}^{n} b_j(y_1)(e^{\alpha x})^{p_j}, \tag{6}$$

the coefficients b_j being real constants with values depending upon y_1. The expression (6) is a polynomial in the positive quantity $e^{\alpha x}$. As such it can vanish, by Descartes's rule, at most as many times as there are changes of sign in the sequence of values b_j. This cannot exceed $(n-1)$, since the number of terms in (6) is n.

Let R designate the rectangular region cut from the strip (5) by two lines $y = y_1$, and $y = y_2$, chosen so that on neither of them there lies a zero of $\Phi(z)$. (There will evidently be such lines through any interval of the y axis, however small.) Then as z traces either of these lines the value arg $\Phi(z)$ can change (increase or decrease) by at most $n\pi$ since $I\{(1/a_0)\Phi(z)\}$ vanishes no oftener than $(n-1)$ times. As z traces the side $x = -K$ of

the rectangle R the value of the sum $\Phi(z)$ remains arbitrarily near that of its first term a_0 (provided K has been chosen sufficiently large), and the change in arg $\Phi(z)$ accordingly remains arbitrarily near to 0. Lastly as z traces the side $x=K$ the change in arg $\Phi(z)$ is an increase in value arbitrarily near to $c_n(y_2-y_1)$, since $\Phi(z)$ on this side of the rectangle is approximated with an arbitrary degree of accuracy by its last term $a_n e^{c_n z}$. From this evaluation of the extreme possible changes in arg $\Phi(z)$ as z traces the boundary of the rectangle it must be concluded that the number $n(R)$ of zeros of $\Phi(z)$ within the rectangle is subject to the bounds

$$(7) \qquad -n+\frac{c_n}{2\pi}(y_2-y_1) \leqq n(R) \leqq n+\frac{c_n}{2\pi}(y_2-y_1).$$

This evidently limits, both above and below, the density of zeros in any portion of the strip (5). In particular it shows that no sum of the form (3) with $(n+1)$ terms can have a zero of multiplicity greater than n.

THEOREM 1. *If in the exponential sum* (1) *the coefficients are constants and the exponents are real and commensurable, then the distribution of zeros is given explicitly by the formula* (4). *In this distribution the number of zeros which lie between two lines* $y=y_1$, *and* $y=y_2$, *is restricted by the relation* (7).

The hypothesis of commensurability of the exponents is evidently a restrictive one. Yet many functions $\Phi(z)$ of importance and interest are included by it. It will be observed, for instance, that the hypothesis is vacuous whenever the sum involves only two terms, and the results obtained apply, therefore, to all such sums. Similarly the familiar trigonometric sums

$$(8) \qquad \begin{aligned} \Phi_c(z) &= \sum_{j=0}^{n} a_j \cos jz, \\ \Phi_s(z) &= \sum_{j=1}^{n} a_j \sin jz \end{aligned}$$

are included, as is in fact any linear combinaton of trigonometric functions of integral multiples of z (iz naturally replacing z). While the problem is in this case theoretically simple, the interrelation of algebraic and transcendental features may be

of distinct interest. The following theorem, due to Pólya,* is at once illustrative of this point and useful for later reference.

THEOREM 2. *If the coefficients a_j are real and the zeros of the polynomial*

$$P(\xi) = \sum_{j=0}^{n} a_j \xi^j \tag{9}$$

all lie within the unit circle about $\xi = 0$, then the zeros of the corresponding trigonometric sums (8) *are all real and simple. Each of these sums has precisely $2n$ zeros on the interval $0 \leqq z < 2\pi$ and the zeros of either sum alternate with those of the other. (By a theorem of Kakeya the hypothesis is fulfilled if $0 \leqq a_0 < a_1 < \cdots < a_n$.)*

The proof is exceedingly simple. Under the hypothesis made arg $P(\xi)$ increases by $2n\pi$ as ξ traces the unit circle, that is, arg $P(e^{iz})$ increases by that amount as z traces the axis of reals from 0 to 2π. But when z is real the sums (8) are respectively the real and imaginary parts of $P(e^{iz})$. Hence these sums vanish in alternation, each $2n$ times on the interval $0 \leqq z < 2\pi$. That the zeros are thus all accounted for follows from the fact that each of the sums (8) may be expressed as a polynomial of degree $2n$ in e^{iz}.

3. *Constant Coefficients and General Real Exponents.* When the exponents are not commensurable the determination of the distribution of zeros of $\Phi(z)$ is not in general of an algebraic character. Independent discussions of the problem in this case have been given by C. E. Wilder[2] and by J. D. Tamarkin.[3] Inasmuch as the subject has to some extent been anticipated by the discussion of the preceding section a slight amplification of the reasoning there used will suffice here for the derivation of such results as have been obtained.

The sum $\Phi(z)$ for the case in hand is expressed by the formula

$$\Phi(z) = \sum_{j=0}^{n} a_j e^{c_j z}, \quad c_0 = 0. \tag{10}$$

Given such a function it is easily seen that with K chosen as a constant sufficiently large the zeros of $\Phi(z)$ all lie within a strip of the z plane determined by a relation (5). This may be de-

* See list of references at the end of this paper.

duced from the dominance of the first term of the sum (10) over all remaining terms for values of z satisfying the condition $x<-K$, and the similar dominance of the last term of the sum when $x>K$. The change in arg $\Phi(z)$ as z traces any portion of the lines $x=-K$ or $x=K$ may be computed precisely as in the case of the preceding section and the establishment of the relation (7) depends only on the consideration of the quantity

$$I\left\{\frac{1}{a_0}\Phi(z)\right\} = \sum_{j=1}^{n} b_j(y)e^{\alpha_j x}, \tag{11}$$

on a line for which y is constant. Since the expression (11) is now not a polynomial Descartes's rule may not be invoked. It is, however, easily proved that in this case also the expressions can vanish at most as many times as there are changes in sign in the sequence of coefficients (see Pólya and Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, p. 49) and hence the relation (7) is valid as before for any rectangle R cut from the strip (5) by a pair of lines $y=y_1$ and $y=y_2$ on neither of which $|\Phi(z)|$ is zero. The existence of such lines through any interval of the y axis follows, of course, from the regularity of the function $\Phi(z)$.

It is a fact of considerable importance in many considerations that in the case of an exponential sum the value $|\Phi(z)|$ is uniformly bounded from zero when the variable z is uniformly bounded from the zeros of $\Phi(z)$. This means that given a sufficiently small positive δ there exists a constant H depending only on δ and such that

$$|\Phi(z)| > H, \text{ for } |z-z_m| > \delta, \tag{12}$$

where z_m designates the set of zeros of $\Phi(z)$. The fact is obvious when consideration is restricted to any finite portion of the z plane. In the case of commensurable exponents it then follows for the entire strip (5) and hence for the entire z plane because of the periodicity of the function $\Phi(z)$. The fact may also be established (see Wilder[2] or Tamarkin[3]), though less simply, in the general case.

THEOREM 3. *If in the exponential sum* (10) *the coefficients are constants and the exponents are real, then the zeros of the sum all lie within a strip* (5), *and in any portion of this strip the number*

of zeros is limited by the relation (7). *When z is uniformly bounded from the zeros of $\Phi(z)$, then $|\Phi(z)|$ is uniformly bounded from zero.*

It may not be devoid of interest to remark that the distance between distinct zeros of $\Phi(z)$ admits of no positive lower bound in the general case of incommensurable exponents. This is shown by the simple and explicit example $\Phi(z)=(1-e^{\alpha z})(1-e^{\beta z})$, in which α/β is irrational. The formula (4) shows that when the exponents are commensurable a minimum distance between zeros does exist.

4. *Coefficients Asymptotically Constant.* In many considerations associated with the distribution of zeros of a function of type (1) the interest centers directly only upon the character of this distribution in the region of the z plane remote from the origin, that is, in the region $|z|>M$, where M is a constant arbitrarily large. When such is the case it is natural that the hypothesis of the constancy of the coefficients may be replaced without essential loss by a less stringent one to the effect that the coefficients be sufficiently like constants when $|z|$ is large. To facilitate the formulation of such hypotheses we shall agree upon the use of symbolism as follows. In a region R of the z plane including the point $z=\infty$ a function will be called an epsilon function provided that it is analytic in every finite portion of R and that it approaches zero uniformly in R as $|z|\to\infty$. To indicate such a function we shall use the generic symbol $\epsilon(z)$.

Let it be supposed now either that the coefficients $A_j(z)$ in the sum (1) are single valued and of the form

$$A_j(z) = a_j + \epsilon(z), \tag{13}$$

in the region $|z|>M$, or that they are multiple valued but such that in the region $|z|>M$, $-\pi<\arg z\leqq\pi$, their various branches are each of the form (13). In the latter case the several determinations of the coefficients correspond to distinct branches of the sum $\Phi(z)$, and the considerations which follow apply directly to any one of these branches. The discussion may be completed by a similar consideration of each individual branch. The form assumed for the sum (1) is, therefore,

$$\Phi(z) = \sum_{j=0}^{n}\{a_j + \epsilon(z)\}e^{\alpha_j z}, \qquad a_0 a_n \neq 0. \tag{14}$$

As already indicated it will be made a part of the hypothesis that neither of the two constants a_0, a_n is zero. It will be shown that the zeros of the sum (14) are asymptotically represented by those of the related sum

$$\Phi_1(z) = \sum_{j=0}^{n} a_j e^{\alpha_j z}, \tag{15}$$

which is of the type considered in §§2 and 3.

It may be shown to begin with, and by precisely the reasoning employed in the analogous consideration of §3, that there can be no zeros of the sum (14) outside a strip of the form (5) when $|z| > M$. Within this strip, on the other hand, each exponential is bounded and hence the function (14) is evidently expressible in the form

$$\Phi(z) = \Phi_1(z) + \epsilon(z), \qquad |x| < K, \ |z| > M.$$

Now in virtue of the distribution of zeros of $\Phi_1(z)$, as determined in the preceding section, these zeros may be enclosed in groups of n or less by a succession of contours γ_r such that no point of such a contour lies within a distance less than δ from a zero of $\Phi_1(z)$, and each of the regions enclosed is of major diameter $n\delta$. For values of z not within such a contour, $|\Phi_1(z)|$ is bounded uniformly from zero, as stated in Theorem 3, and hence for such values of z,

$$\Phi(z) = \Phi_1(z)\{1 + \epsilon(z)\}.$$

From this relation it follows that when $|z| > M$, $\Phi(z)$ has within each contour γ_r precisely as many zeros as $\Phi_1(z)$, and that all zeros of $\Phi(z)$ are thereby accounted for.

THEOREM 4. *If the function $\Phi(z)$ (or a determination of it) is of the form* (14), *then in the region $|z| > M$ the distribution of zeros of $\Phi(z)$ (or of the branch of $\Phi(z)$ in question) may be described as in Theorem* 3. *The zeros are asymptotically represented by those of the related sum* (15).

5. *Coefficients which are Asymptotically Power Functions.* It will be supposed now that in the form (1) the coefficients $A_j(z)$, or chosen branches of them, are of the form

$$A_j(z) = z^{\nu_j}\{a_j + \epsilon(z)\}, \qquad a_0 a_n \neq 0, \tag{16}$$

for z in the region $|z|>M$, $-\pi<\arg z\leqq\pi$. The values ν_j are taken to be any real constants positive, negative or zero.

The sums with coefficients of type (16) evidently include as a sub-class those with coefficients which are polynomials in z. A discussion of the latter was given first by Pólya[6] and Schwengler.[7] Their method, which is largely geometric, is in many respects convenient and illuminating and will be utilized in the following discussion whenever it is of advantage.

Without further convention the values ν_j (except for $j=0$, and $j=n$) are not fixed beyond ambiguity by the form (16), inasmuch as an increase in ν_j may be compensated for by the choice of a_j as zero. It will be assumed, therefore, for the sake of definiteness that whenever the form of the coefficient $A_j(z)$ permits, the value ν_j will be chosen so that the corresponding constant a_j is not zero. The terms of the sum for which this is possible will be called ordinary terms. In the exceptional non-ordinary terms the value of ν_j must remain unspecified. Under the form (16) the constant a_j for every such term will be zero. Thus, by way of example, if $A_j(z)=\log z$, then the form (16) implies $a_j=0$, with $\nu_j>0$.

As a preliminary to the discussion of the general case under (16) it will be convenient to consider that in which all terms are ordinary terms, and in which the constants ν_j are proportional to the exponents c_j.

6. *The Values ν_j and c_j Proportional.* If the real constant β is defined by the relation

$$\nu_j = \beta c_j, \qquad (j = 1, 2, \cdots, n),$$

the formula

$$\zeta = z + \beta \log z \tag{17}$$

defines a single-valued analytic map of the portion of the z plane to which z was restricted above upon a complex ζ plane, the point $z=\infty$ corresponding to $\zeta=\infty$. It follows that any epsilon function of z is also such a function of ζ, and it is evident, therefore, that under the relation (17)

$$\Phi(z) \equiv \sum_{j=0}^{n} \{a_j + \epsilon(\zeta)\} e^{c_j \zeta}, \qquad a_0 a_n \neq 0.$$

This, however, as a function of ζ is precisely of the form (14), and in the ζ plane the zeros are accordingly distributed as described in Theorem 4. It is necessary only to translate this result into terms of the variable z and this can be readily done. Thus if $\zeta=\xi+i\eta$ the components of the relation (17) are

$$\xi = x + \beta \log|z|,$$

$$\eta = y + \beta \arg z.$$

The rectilinear strip $|\xi|<K$ corresponds, therefore, to the curvilinear strip bounded by the logarithmic curves

$$x + \beta \log|z| = \pm K. \tag{18}$$

This strip which is asymptotically of constant width $2K$, and whose bounding curves approach parallelism with the axis of imaginaries, contains all the zeros of $\Phi(z)$ in the region $|z|>M$. Since for z in the strip (18) the value arg z approaches either $\pi/2$ or $-\pi/2$ according as the upper or lower half plane is considered, it follows that any line $\eta=\eta_1$ on which η is constant maps asymptotically upon the line $y=\eta_1\pm\beta\pi/2$, on which y is constant. Moreover, the distance between two such lines both in the upper or both in the lower half plane is asymptotically preserved.

THEOREM 5. *If in the exponential sum* (1) *the coefficients are of the form* (16) *with values ν_j proportional to the exponents c_j, and all terms are ordinary terms, then the zeros of the sum are asymptotically located within a logarithmic curvilinear strip bounded by curves of the form* (18), *and the number of zeros lying between any two lines parallel to the axis of reals is asymptotically subject to the relations* (7).

7. *General Real Values ν_j.* With each ordinary term of the sum

$$\Phi(z) = \sum_{j=0}^{n} z^{\nu_j}\{a_j + \epsilon(z)\}e^{c_j z}, \tag{19}$$

there can be associated in the z plane the corresponding point P_j with Cartesian coordinates (c_j, ν_j). These points determine the broken (or straight) line L, which (i) joins P_0 with P_n, (ii) has vertices only at points of the set P_j, (iii) is convex upward (or straight), and (iv) is such that no points of the set lie above it. Let L_r designate the rth segment of this broken line L, the

order being from left to right, and let the scheme of single subscript notation employed in (19) and in designating the set of points P_j be replaced by one in which the points of the set lying on the segment L_r are denoted in order by P_{rh}, $(h=1, 2, \cdots, n_r)$.

If the sum (19) contains exceptional terms it must be made a matter of hypothesis that for every such term a choice of the value ν_j is possible under the form (16) such that the corresponding point (c_j, ν_j) lies below the broken line L. All points of the set located below the line L may then be designated in any order by P_{0h}, $(h=1, 2, \cdots, n_0)$, similar double subscripts being assigned to the corresponding terms in (19).

Consider now any intermediate segment L_r of the line L, and let its slope be m_r, the slope of the preceding segment being m_{r-1}. Then as the real parameter k varies over the range

$$m_{r-1} - \epsilon \geqq k > m_r - \epsilon,$$

with ϵ positive and sufficiently small but otherwise arbitrary, the curve

$$x = -k \log |z| \tag{20}$$

sweeps out the region of the z plane bounded by the curves

$$\begin{aligned} x &= -(m_{r-1} - \epsilon) \log |z|, \\ x &= -(m_r - \epsilon) \log |z|. \end{aligned} \tag{21}$$

The figure is symmetrical in the upper and lower half planes. On any curve (20) it is obvious that

$$|z^{\nu_{ls}} e^{c_{ls} z}| = |z|^{\nu_{ls} - k c_{ls}},$$

and hence the relative magnitude of the terms in (19) for z on this curve depends upon the corresponding values $(\nu_{ls} - k c_{ls})$. These, however, admit of ready geometrical interpretation. They are, namely, the y intercepts of the lines having the slope k and passing through the respective points P_{ls}. In virtue of this fact and the restriction upon the range of k it is readily seen that for every $l \neq r$ the inequality

$$(\nu_{ls} - k c_{ls}) < (\nu_{r1} - k c_{r1})$$

obtains. It follows, therefore, that the form of expression

$$z^{\nu_{ls}} e^{c_{ls} z} = z^{\nu_{r1}} e^{c_{r1} z} \epsilon(z), \quad l \neq r,$$

is valid for z on any curve (20), and hence that the sum (19) can be written in the form

$$\Phi(z) = \sum_{h=1}^{n_r} z^{\nu_{rh}} \{a_{rh} + \epsilon(z)\} e^{c_{rh}z}, \tag{22}$$

for all z of the region bounded by the curves (21). Upon removal of a factor $z^{\nu_{r1}}e^{c_{r1}z}$, however, the sum (22) takes the form considered in §6, since the constants $(\nu_{rh}-\nu_{r1})$ are proportional to the exponents $(c_{rh}-c_{r1})$, the constant of proportionality being m_r, the slope of the segment L_r. It may be concluded, therefore, by Theorem 5, that the zeros of $\Phi(z)$ in the region (21) are asymptotically confined to the logarithmic strip bounded by the curves

$$x + m_r \log|z| = \pm K, \tag{23}$$

and that the number of zeros in this strip and between two lines $y=y_1$ and $y=y_2$ is restricted by the relation obtained from (7) by replacing n and c_r by n_r and $(c_{r,n_r}-c_{r1})$.

The geometric picture is completed if it is observed that in the upper half of the plane the central line of the strip (23), that is, the curve

$$x + m_r \log|z| = 0, \tag{24}$$

passes through the point (0, 1) with slope $-1/m_r$, namely in the direction perpendicular to that of the associated segment L_r. There is a strip of the form (23) associated with each segment of the broken line L (obvious simple modifications of the discussion being necessary in the case of an end segment), and the zeros of $\Phi(z)$ in the remote part of the z plane are all contained within these strips.

THEOREM 6. *If $\Phi(z)$ is an exponential sum with coefficients of the form* (16) *(the ν_j for exceptional terms satisfying the hypothesis of the text), then the zeros of the sum are asymptotically confined to a finite number of logarithmic strips* (23), *the number of zeros in any strip between two lines parallel to the axis of reals being asymptotically subject to a relation similar to* (7).

8. *Collinear Complex Exponents.* The hypotheses which have been made in the various cases of the sum (1) hitherto considered have been concerned primarily with the structure of the

coefficient functions $A_j(z)$. The exponents have been assumed throughout to be real. It will be shown now that this assumption is dispensable, and that the distribution of zeros of the sum having complex exponents is also determinable. Preliminary to the general discussion the considerations of this section will be given to the case in which the values c_j viewed as points in the complex plane are collinear.

Let l designate the line on which the points c_j are located, and let the subscripts be assigned to these points in the order of their location on l. Then if θ designates the inclination angle of the line l with respect to the axis of reals (the positive sense of l being from c_0 to c_n), it follows that the relations

$$c_j = c_0 + \gamma_j e^{i\theta}, \qquad (j = 0, 1, 2, \cdots, n),$$

are satisfied by a set of real values $0 = \gamma_0 < \gamma_1 < \cdots < \gamma_n$. In terms of the variable ζ defined by the relation

$$\zeta = ze^{i\theta}, \tag{25}$$

the sum (1), therefore, takes the form

$$\Phi = e^{c_0 z} \sum_{j=0}^{n} B_j(\zeta) e^{\gamma_j \zeta}, \tag{26}$$

where $B_j(\zeta) \equiv A_j(z)$ under the substitution (25). The sum in (26) is one in which the exponents γ_j are real, and the coefficients $B_j(\zeta)$ have the essential structural characteristics of $A_j(z)$. The distribution of zeros of the sum Φ in the ζ plane is, therefore, as described in the theorems of the preceding sections, that is, the zeros are asymptotically confined to one or more strips which are parallel or approach parallelism with the axis $\arg \zeta = \pm \pi/2$. Since the transformation (25) is a mere rotation of the plane, it may be concluded that in the z plane the zeros of the sum are confined for $|z| > M$ to a similar set of strips parallel or approaching parallelism with a direction $\arg z = -\theta \pm \pi/2$. These directions, it will be noted, are those perpendicular to the line in the complex plane which contains the points $\bar{c}_j$, the complex conjugates of the exponents c_j.

THEOREM 7. *If the exponents c_j in the exponential sum* (1) *are collinear complex constants, the distribution of zeros of $\Phi(z)$ is obtainable from the theorem previously enunciated by substituting*

in the role of the axis of reals the line containing the points $\bar{c}_j$ conjugate to the exponents c_j.

9. *General Complex Constants.* In the most general type of exponential sum to which consideration will here be given the exponents may be any set of complex constants. This case is now mature for discussion. While treatments of it by purely analytic methods are available[2,3,4,5] and effective, a geometric method[6] will again be resorted to as convenient for the purpose at hand. Let it be supposed that the values $\bar{c}_j$ (conjugate to the exponents) have been plotted in the complex plane, and let the points of the set so represented be surrounded by the polygon P which (i), is convex, (ii), has vertices only at points of the set, and (iii), includes all points of the set either in its interior or on its perimeter.

Let the sides of the polygon P taken in counterclockwise succession with any one as the initial one be designated by l_r, $r=1, 2, \cdots, q$, and with this done let the points of the set $\bar{c}_j$ which lie on the side l_r be redesignated by $\bar{c}_{rh}$, $h=1, 2, \cdots, h_r$, the succession being again in the counterclockwise sense around the polygon P. The points $\bar{c}_j$ lying within the polygon may in conclusion be denoted (in any order) by $\bar{c}_{0h}$, $h=1, 2, \cdots, h_0$.

If ω_r represents the inclination angle of the outer normal to the side l_r, the relation

$$\omega_{r-1} + \epsilon \leqq \tau < \omega_r + \epsilon,$$

with ϵ sufficiently small and positive but otherwise arbitrary, restricts the real parameter τ. As τ varies over the range permitted, the line

$$\arg z = \tau \tag{27}$$

sweeps out in the z plane a sector described by the relation

$$\omega_{r-1} + \epsilon \leqq \arg z < \omega_r + \epsilon. \tag{28}$$

Such a sector is thus associated with each side of the polygon, and the set of these sectors completely fills out the complex plane. It is proposed to consider the sum $\Phi(z)$ within the typical sector (28). The extension of the results to the entire plane will then be immediate.

When z lies on a line (27) it is readily computed that

$$\left| z^{\nu_{sk}} e^{c_{sk} z} \right| = e^{|z|\{\alpha_{sk}\cos\tau+\beta_{sk}\sin\tau+\nu_{sk}\log|z|/|z|\}},$$

where $\bar{c}_{sk}=\alpha_{sk}+i\beta_{sk}$. It follows from this that the relative magnitudes of the terms in the sum (1) are ultimately determined by the relative magnitudes of the associated quantities

$$\alpha_{sk}\cos\tau+\beta_{sk}\sin\tau.$$

These, however, are represented geometrically by the respective projection of the vectors $\bar{c}_{sk}$ upon the line (27), and from this interpretation it is quickly seen, in virtue of the restriction upon the range of τ, that

$$\alpha_{sk}\cos\tau+\beta_{sk}\sin\tau<\alpha_{r1}\cos\tau+\beta_{r1}\sin\tau, \text{ for any } s\neq r.$$

It follows that for z on any line (27), that is, anywhere in the sector (28), the form of expression

$$z^{\nu_{sk}} e^{c_{sk} z} = z^{\nu_{r1}} e^{c_{r1} z}\epsilon(z)$$

is valid for $s\neq r$ and any k. The sum $\Phi(z)$ with coefficients of the form (16) may, therefore, be expressed as a sum with a reduced number of terms, that is,

$$\Phi(z) = \sum_{k=1}^{h_r} A_{rk}(z) e^{c_{rk} z}, \tag{29}$$

and in this expression which is valid for all z in the section (28) the coefficients are still of the form (16). The exponents in the sum (29) are, however, collinear, their conjugates $\bar{c}_{rk}$ lying on the side l_r. It may, therefore, be concluded from Theorem 7 that the zeros of $\Phi(z)$ within the sector (28) are asymptotically confined to one or more strips which are parallel or approach parallelism with the normal to the associated side of the polygon. A similar conclusion may obviously be drawn in connection with the values of z in any other sector.

THEOREM 8. *If in the sum* (1) *the exponents are any complex constants, the zeros of* $\Phi(z)$ *are confined for* $|z|>M$ *to a finite number of strips each of asymptotically constant width. These strips are associated in groups with the exterior normals to the sides of the polygon described in the text, and approach parallelism with the respective normals. Within each group of strips the dis-*

tribution of zeros may be described as in the previously stated theorems, the role of the axis of reals being transferred to the respective side of the polygon.

In closing here this brief and necessarily incomplete discussion of the distribution of the zeros of an exponential sum, I wish at least to call attention to the related work which has been done by Ritt[8] on the factorization of such sums and on algebraic combinations of exponentials.

Part II. The Exponential Integral

10. *Discussion of the Integral.* If in the exponential sum with constant coefficients and real exponents as given in form (9), the origin is assumed to be a zero of the sum (and this can be made so by a simple translation of the plane), then $\sum_{j=0}^{n} a_j = 0$. With this adjustment the sum $\Phi(z)$ can be simply represented in the form of an integral as follows. Let the coefficient function $\phi(t)$ be constant in successive intervals as defined by the relation

$$\phi(t) \equiv -\sum_{j=0}^{r} a_j,$$

for $c_r \leqq t < c_{r+1}$, $(r = 0, 1, 2, \cdots, n-1)$. Then it is found directly that

$$\Phi(z) = z\int_{c_0}^{c_n} \phi(t)e^{zt}dt.$$

The exponential integral of the general type

$$\Psi_{a,b}(z) = \int_a^b \psi(t)e^{zt}dt \tag{30}$$

thus includes among its specializations (except for a factor z) the exponential sums of the types considered in §2 and §3, and hence, of course, also all trigonometric functions which are so expressible. Besides these a number of other functions, including Bessel functions, are amenable to representation in this form.

In turning to a consideration of the integral (30) a restriction of the discussion is necessitated by the brevity of available time. It is proposed, therefore, to summarize only certain investiga-

tions which deal with the integral with finite limits of integration and which impose none but qualitative hypotheses on the coefficient function $\phi(t)$ involved. Many investigators have considered either integrals of specialized forms included under the type (30), or the integral with infinite limits. It is not due to a want of interest in the results obtained that these subjects are omitted from the present discussion.

Various normalizations of the integral (30) may be attained by simple substitutions on the variables involved. Appropriate normalizations of this kind make for simplicity in the enunciation of hypotheses and results, and for this reason the following ones will be assumed as indicated from point to point in the ensuing discussion. To begin with, the substitution of $\alpha t+\beta$ and z/α, respectively, for t and z affects a change in the limits of integration, and by suitable choice of the constants α, β, any desired finite limits may be attained. It will be found of particular convenience to normalize the interval thus in some cases to (0, 1) and in others to (−1, 1). Further, the substitution of $z+c$ for z affects the introduction of a factor e^{ct} into the coefficient function $\psi(t)$, and it follows that whenever $\psi(a)\neq 0$ and $\psi(b)\neq 0$, a normalization may be made such that $\psi(a)=\psi(b)=1$. Lastly it will be observed that the trigonometric integrals

$$\text{(31)} \qquad \text{(a)} \qquad \Psi_s(z) = \int_0^1 f(t) \sin zt \, dt,$$

$$\text{(b)} \qquad \Psi_c(z) = \int_0^1 f(t) \cos zt \, dt$$

are both expressible in the form $\Psi_{-1,1}(iz)$, the coefficients being given respectively by the formulas $[1/(2i)]\{f(t)-f(-t)\}$, and $(1/2)\{f(t)+f(-t)\}$.

11. *The Integral with an Infinity of Real Zeros.* The distribution of zeros of an exponential sum (9) as determined in § 3 features conspicuously the existence of a bound upon the numerical values of x for which a zero of the sum may occur. This fact, it will be recalled, followed from the ultimate dominance of the one or the other extreme term of the sum when $|x|$ becomes sufficiently large. The integral (30) is in one sense the analogue of an exponential sum with infinitely many exponents lying on the range between a and b. Either of these extreme

values might in such event be a limit point for the infinite set, and if such were the case it would conceivably involve an indefinite delay in the accession to dominance of the respective extreme term of the sum. The absence of the corresponding bound upon the x coordinate of the zeros might then be reasonably expected. That the situation thus conceived does as a matter of fact exist is shown by the following example constructed by Titchmarsh.[9]

Let the constants θ and ν be chosen subject to the conditions $\theta<1/2$, $\nu>2$, and let $\alpha=\nu/\theta$. The function

$$\psi(t) = \begin{cases} (-\alpha)^k e^{-\nu^k}, & \text{for } \theta^k - \alpha^{-k} < t \leqq \theta^k \quad (k = 1, 2, 3, \cdots), \\ 0, & \text{elsewhere on } (0, 1), \end{cases}$$

is integrable (absolutely) in the sense of Lebesgue. But for any positive integer n it is found respectively that

$$\left| \int_0^{\theta^{n+1}} \psi(t) e^{-\alpha^n t} dt \right| \leqq \int_0^{\theta^{n+1}} |\psi(t)|\, dt$$

$$= \sum_{k=n+1}^{\infty} e^{-\nu^k} < e^{-\nu^{n+1}} \sum_{j=0}^{\infty} e^{-\nu^j} = e^{-\nu^{n+1}} \int_0^1 |\psi(t)|\, dt;$$

that

$$\left| \int_{\theta^{n-1}-\alpha^{-n+1}}^{1} \psi(t) e^{-\alpha^n t} dt \right| \leqq e^{-\alpha^n(\theta^{n-1}-\alpha^{-n+1})} \int_0^1 |\psi(t)|\, dt$$

$$= e^{-\nu^n/\theta+\alpha} \int_0^1 |\psi(t)|\, dt;$$

and that

$$\int_{\theta^n-\alpha^{-n}}^{\theta^n} \psi(t) e^{-\alpha^n t} dt = (-1)^n (e-1) e^{-2\nu^n}.$$

Hence it follows that

$$\left| e^{2\nu^n} \Psi_{01}(-\alpha^n) - (e-1)(-1)^n \right|$$

$$< \left\{ e^{\nu^n(2-\nu)} + e^{\nu^n(2-1/\theta)+\alpha} \right\} \int_0^1 |\psi(t)|\, dt,$$

and since the quantity on the right approaches zero with $1/n$ it must be concluded that $\Psi_{0,1}(z)$ takes when $z=-\alpha^n$ (and n is sufficiently large) the sign of $(-1)^n$. The existence of at least one zero in each corresponding interval $-\alpha^{n+1}\leqq z\leqq -\alpha^n$ is thus established.

12. *Hypotheses of Specific Type.* While the example of the preceding section shows that the distribution of zeros of an exponential integral may differ sharply from that of an exponential sum, yet in a very wide class of cases they show considerable similarity. The coefficient function $\psi(t)$ is of course the determinative factor of the integral, and hence is always the direct subject of the hypothesis. The substance of the present section is essentially an abstract of results which are due to Pólya[1] and is characterized by the fact that the hypotheses are on the whole of rather specific character. The results are accordingly precise to a considerable degree of detail.

If the function $\psi(t)$ is a step-function with discontinuities finite in number and occurring only at rational points of the interval, the function will be called exceptional. With such a function as coefficient the integral may be directly evaluated into the form of an exponential sum of the elementary type considered in §2. Hence no essential loss is involved in the restriction of the theorems to unexceptional functions, and this will be tacitly assumed in the absence of statements to the contrary.

THEOREM 9. *If the function $\psi(t)$ is positive and non-decreasing on the interval (0, 1), then the zeros of the integral $\Psi_{0,1}(z)$ are all located in the open half-plane $x<0$.*

By way of proof let the value $z=iy$ be first considered. For any value t_1 on the interval of integration it is seen that

$$\left| iy\int_0^{t_1}\psi(t)e^{iyt}dt\right| \geqq \psi(t_1)-\psi(0)-\left|\int_0^{t_1}e^{iyt}d\psi\right|.$$

But if θ is defined by the relation

$$\left|\int_0^{t_1}e^{iyt}d\psi\right| = e^{-i\theta}\int_0^{t_1}e^{iyt}d\psi,$$

then as $t_1\to 1$ the relation above becomes

$$| iy\Psi_{01}(iy) | \geqq \lim_{t_1 \to 1} \int_0^{t_1} \{1 - \cos(yt - \theta)\} d\psi.$$

Since $d\psi$ is never negative, the conclusion $|\Psi_{01}(iy)| > 0$ will be expected, and is in fact correct. Its derivation is, however, impeded somewhat by the necessity of recognizing and disposing of the potential effect of the functional dependence of θ upon the variable t_1. In conclusion the substitution of $z = x + iy$ in the discussion given merely replaces $\psi(t)$ by $\psi(t)e^{xt}$, and since this satisfies the hypotheses made for any fixed non-negative x, the assertion of the theorem is established.

The result obtained admits of ready extension into the following form.

THEOREM 10. *If the coefficient $\psi(t)$ is positive and continuous, and if $\psi'(t)$ exists (except possibly at a finite number of points) and*

$$\alpha \leqq -\frac{\psi'(t)}{\psi(t)} \leqq \beta,$$

then the zeros of the integral $\Psi_{0,1}(z)$ are confined to the open strip of the complex plane given by the relation

$$\alpha < x < \beta.$$

(The trivial case $\psi(t) = e^{ct+d}$ is excepted.)

The transformations of the integral into the forms

$$\Psi_{0,1}(z) = \int_0^1 e^{\beta t}\psi(t)e^{(z-\beta)t}dt = e^z \int_0^1 e^{-\alpha t}\psi(1-t)e^{(\alpha - z)t}dt$$

are easily verified. In these forms the coefficients $e^{\beta t}\psi(t)$, and $e^{-\alpha t}\psi(1-t)$, satisfy the hypothesis of Theorem 9, and hence the zeros are simultaneously restricted by the relations

$$R(z - \beta) < 0, \quad R(\alpha - z) < 0.$$

From this the theorem follows.

THEOREM 11. *If the function $f(t)$ is positive and non-decreasing on the interval (0, 1), then the zeros of the trigonometric integrals $\Psi_s(z)$ and $\Psi_c(z)$ defined in (31) are all real and simple. The zeros of $\Psi_c(z)$ occur singly in each of the intervals*

$$[(m - 1/2)\pi, (m + 1/2)\pi], \qquad m \neq 0,$$

while those of $\Psi_s(z)$ *occur (except for* $z=0$*) singly in each of the intervals* $[m\pi, (m+1)\pi]$, $m \neq 0, -1$.

The facts of this theorem may be established by suitable passages from finite to corresponding infinite sums as is indicated in the following. The trigonometric sum

$$\Phi_s(z) = \frac{1}{n} \sum_{j=0}^{n} f\left(\frac{j}{n}\right) e^{ij/n^2} \sin\left(\frac{j}{n} z\right)$$

has, for each value of n, coefficients which are positive and increasing. Its zeros, therefore, by Theorem 2, are real, and alternate with those of the corresponding sum of cosines. From this it may be concluded, with suitable reasoning as $n \to \infty$, that the zeros of the integrals $\Psi_s(z)$ and $\Psi_c(z)$ are real, and that a multiple zero of the one would necessarily be also a zero of the other. Since in this latter event the integral

$$\Psi_c(z) + i\Psi_s(z) = \int_0^1 f(t) e^{izt} dt$$

would have a real zero, contrary to Theorem 9, the simplicity of the zeros must be granted.

The more precise location of the zeros as stated in the theorem is derived by Pólya from the expansion

$$(32) \quad \frac{\Psi_s(z)}{z \sin z} = \lim_{n \to \infty} \left\{ \frac{\Psi_s'(0)}{z} - \sum_{j=1}^{n} \frac{(-1)^{j+1} \Psi_s(j\pi)}{j\pi} \left[\frac{1}{z - j\pi} + \frac{1}{z + j\pi} \right] \right\},$$

which may be established by methods familiar in the theory of functions of a complex variable. Each coefficient on the right of this expansion is positive, for

$$(-1)^{j+1} \Psi_s(j\pi) = (-1)^{j+1} \int_0^1 f(t) \sin j\pi t dt,$$

while the integral represents a sum of areas, alternating in sign, non-decreasing in numerical magnitude, with the final one positive. Because of this the rational function within the brace in (32) is found by elementary means to have one zero in each of the intervals $(-\infty, -n\pi)$, $(n\pi, \infty)$, $[m\pi, (m+1)\pi]$, $m = -n, \cdots, -1, 1, 2, \cdots, n-1$. A suitable passage to the limit as

$n \to \infty$ establishes the theorem. An analogous discussion for $\Psi_c(z)$ is, of course, possible.

If the hypothesis of Theorem 11 is extended to include the assumptions that $\lim_{t\to 0} f(t) = 0$, and that $f(t)$ is nowhere concave, a refinement of the theorem is possible which determines the specific half of each of the intervals given in which the zero is located. The method of proof is similar to that outlined above, a remark which applies also to the following theorem.

THEOREM 12. *If the function $f(t)$ is positive, decreasing, and nowhere convex downward, and if $f'(t)$ exists and is not an exceptional function, then the zeros of $\Psi_c(z)$ are all real and there is just one in each of the intervals $[m\pi, (m+1)\pi]$, $m \neq 0, -1$.*

As examples of integrals covered by the theorems of this section one has the familiar functions

$$J_0(z) = \int_0^1 \frac{2}{\pi(1-t^2)^{1/2}} \cos zt\,dt,$$

$$J_1(z) = \int_0^1 \frac{2t}{\pi(1-t^2)^{1/2}} \sin zt\,dt.$$

13. *Hypotheses of Intermediate Type.* Less specific than those of the preceding section are the hypotheses and results to which the attention will now be turned. The material has been taken essentially from a memoir by Miss M. L. Cartwright[10] under date of the present year. The two theorems to be given first, however, are attributed by her to Hardy.

As a preliminary it is perhaps desirable to note at this point the following relations which have been proved by Titchmarsh,[9] namely that with any coefficient function $\psi(t)$ which is integrable,

$$(33) \qquad |\Psi_{a,b}(z)| \begin{cases} = o(e^{|z|}), \\ \neq O(e^{|z|-|z|^\delta}), \text{ for any } \delta > 0. \end{cases}$$

These assertions may be established by a direct analysis of the integral involved.

THEOREM 13. *If the function $\psi(t)$ is itself an integral, if it is continuous at $t=1$ and $t=-1$, and if $\psi(1)=\psi(-1)=1$, then the*

zeros, z_m, of the integral $\Psi_{-1,1}(z)$ are given by the asymptotic formula

$$z_m \sim m\pi i.$$

The proof is immediate. An integration by parts yields the equality

$$z\Psi_{-1,1}(z) = e^z - e^{-z} - \int_{-1}^{1} \psi'(t)e^{zt}dt.$$

Of this the right hand member is of the form

$$e^z\{1 + \epsilon(z)\} - e^{-z}\{1 + \epsilon(z)\},$$

because of the first of the relations (33). That the zeros are located as stated in the theorem may be concluded then from the deductions of §2 and §4.

THEOREM 14. *If the function $\psi(t)$ is of bounded variation, is continuous at $t=1$ and $t=-1$, and $\psi(1)=\psi(-1)=1$, then the zeros of $\Psi_{-1,1}(z)$ all lie within a strip of the complex plane given by a relation $|x|<K$.*

The proof of this theorem has been constructed on the following lines. When $x>0$, an integration yields the evaluation

$$z\Psi_{-1,1}(z) = e^z - e^{-z} - \int_{-1}^{1-\delta} e^{zt}d\psi - \int_{1-\delta}^{1} e^{zt}d\psi,$$

in which the integrals on the right are numerically less, respectively, than $He^{z(1-\delta)}$, and $e^z\eta(\delta)$, with H some constant and $\eta(\delta)$ the maximum of $|\psi(t_2)-\psi(t_1)|$ for $1-\delta<t_1<t_2\leqq 1$. Hence

$$|z\Psi_{-1,1}(z)| \geqq e^z\{1 - e^{-2z} - He^{-\delta z} - \eta(\delta)\}.$$

However, the right member of this inequality is assuredly positive if δ is chosen sufficiently small (because of the continuity of $\psi(t)$ at $t=1$ this makes $\eta(\delta)$ small), and if K is then chosen sufficiently large and $x>K$. It evidently follows that for such values of z the integral can have no zero, and a similar discussion establishes the fact for $x<-K$.

The hypotheses of Theorem 14 are seen from the conclusions thus drawn to be sufficient to assure for the exponential integral a distribution of zeros resembling in at least one salient

feature the distribution associated with an exponential sum. By the following theorem this similarity is shown to extend beyond the mere restriction of the zeros to a strip of the plane even to the density of zeros within that strip.

THEOREM 15. *Under the hypotheses of the preceding theorem the number, $n(r)$, of zeros of the integral $\Psi_{-1,1}(z)$ which are in absolute value less than r satisfies the relation*

$$n(r) = \frac{2r}{\pi} + O(1).$$

It is readily found that this relation is included in (7) which restricts the zeros of exponential sums.

The proofs of Theorem 15, and of the following Theorem 16, are less simple than any which have hitherto been outlined. They are based upon facts and relations familiar in the theory of entire functions and involve considerable detail. The author of the theorem[10] has deduced also a formula for the value of the quantity

$$\int_0^r \frac{n(s)}{s} ds.$$

THEOREM 16. *If the function $\psi(t)$ is continuous, with modulus of continuity $\omega(\delta)$ on the interval $(-1, 1)$, and if $\psi(1) = \psi(-1) = 1$, then the zeros of the integral $\Psi_{-1,1}(z)$ all lie within a region*

$$|x| < K|z|\,\omega\left(\frac{1}{|z|}\right),$$

with a suitably chosen constant K, and

$$n(r) = \frac{2r}{\pi} + O\left(r\omega\left(\frac{1}{r}\right)\right).$$

It will be recalled that the modulus of continuity $\omega(\delta)$ is defined as the maximum value of $|\psi(t_2) - \psi(t_1)|$ for t_2, t_1 on $(-1, 1)$ and $|t_2 - t_1| < \delta$. As $\delta \to 0$, $\omega(\delta) = o(1)$, and in fact if the function satisfies a Lipschitz condition $|\psi(t_2) - \psi(t_1)| < A|t_2 - t_1|^{\alpha}$,

$0<\alpha\leqq 1$, then $\omega(\delta)=O(\delta^{\alpha})$. As in the case of the preceding theorem an evaluation of the quantity

$$\int_0^r \frac{n(s)}{s}ds$$

is also deducible. Moreover, it can be shown that the requirement of continuity over the entire interval $(-1, 1)$ can, with suitable modification in the results, be replaced by hypotheses demanding continuity only near the end points.

14. *The Integral with Most General Coefficient.* The exponential integral of greatest generality results from the admission of any function which is absolutely integrable to the role of the coefficient $\psi(t)$. As will be indicated below even this minimum of hypotheses is not unproductive of significant results.

The function $\Psi_{0,1}(z)$ is analytic over the entire plane, and as such its zeros may be enumerated in the order of increasing absolute magnitude, thus, $z_m=r_m e^{i\theta_m}$. The following theorems involve a qualitative characterization of the density with which these zeros are distributed in the complex plane.

THEOREM 17. *If the coefficient function $\psi(t)$ is absolutely integrable, the zeros z_m of the integral $\Psi_{0,1}(z)$ are distributed so that the series*

$$\sum_{m=1}^{\infty}\frac{1}{|z_m|^{1+\epsilon}}$$

converges for every $\epsilon>0$, and diverges for $\epsilon=0$.

As an entire function the integral $\Psi_{0,1}(z)$ is of order at most unity, as may be seen from the evident relation

$$|\Psi_{0,1}(z)|\leqq e^{|z|}\int_0^1|\psi(t)|\,dt.$$

Inasmuch as this order at least equals the exponent of convergence, the statement of the theorem in so far as it concerns the convergence for $\epsilon>0$ is established. The divergence for $\epsilon=0$ is not so immediately verified. Independent proofs of it have, however, been given by both Pólya[1] and Titchmarsh.[9] The method employed by the latter may be indicated briefly as follows.

Convergence of the series for $\epsilon=0$ would imply the absolute convergence of the canonical product

$$P(z) = \Psi_{01}(0) \prod_{j=1}^{\infty} \left(1 - \frac{z}{z_j}\right).$$

In such event, however, a product of this type is known to satisfy the relation

$$| P(z) | = O(e^{|z|^{\eta}}),$$

with any positive constant η. Since $\Psi_{0,1}(z)$ is expressible in the form $\Psi_{0,1}(z) = e^{(\alpha+i\beta)z}P(z)$, its order being unity, it would follow that

$$| \Psi_{0,1}(\pm x) | = O(e^{(\pm\alpha+\eta)|x|}),$$

and this contradicts the second of the relations (33).

THEOREM 18. *If the function $\psi(t)$ is absolutely integrable, then the zeros of the integral $\Psi_{0,1}(z)$ satisfy the relations*

(a) $$n(r) \sim \frac{r}{\pi},$$

(b) $$\sum_{m=1}^{\infty} \frac{\cos \theta_m}{|z_m|} \quad \textit{converges absolutely},$$

(c) $$\sum_{m=1}^{\infty} \frac{\sin \theta_m}{|z_m|} \quad \textit{converges conditionally}.$$

These results have all been obtained by Titchmarsh[9] by the use of relations in the theory of entire functions. The result (a) evidently implies the assertions of Theorem 17, and is in fact more precise. The result (b) in conjunction with (a) indicates that in general $|\cos \theta_m|$ is small when m is large. The example of §11 would show, however, that a relation $\lim_{m\to\infty} \cos \theta_m = 0$ is not in general valid. The result (c), which is obvious in the case of a real coefficient function $\psi(t)$, since the zeros present themselves then in conjugate pairs, is not easily established in the general case. The author of the result[9] has, however, succeeded in explicitly summing both the series (b) and (c) in terms of the values taken by $\Psi_{0,1}(z)$ and its derivative at $z=0$.

LIST OF REFERENCES

1. Pólya, G., *Über die Nullstellen gewisser ganzer Funktionen*, Mathematische Zeitschrift, vol. 2 (1918), pp. 352–383.

2. Wilder, C. E., *Expansion problems of ordinary linear differential equations, etc.*, Transactions of this Society, vol. 18 (1917), pp. 415–442.

3. Tamarkin, J. D., *Some general problems in the theory of ordinary linear differential equations*, 1917, Petrograd; and Mathematische Zeitschrift, vol. 27 (1927), p. 24.

4. Birkhoff, G. D., *Boundary value and expansion problems, etc.*, Transactions of this Society, vol. 9 (1908), p. 373.

5. Langer, R. E., *The asymptotic location of the roots of a certain transcendental equation*, Transactions of this Society, vol. 31 (1929), p.837.

6. Pólya, G., *Geometrisches über die Verteilung der Nullstellen, etc.*, Sitzungsberichte der Bayerischen Akademie, München, 1920, p. 285.

7. Schwengler, E., *Geometrisches über die Verteilung der Nullstellen, etc.*, Dissertation, Zürich, 1925.

8. Ritt, J. F., *A factorization theory, etc.*, Transactions of this Society, vol. 29 (1927), p. 584.

On the zeros of exponential polynomials, and algebraic combinations of exponentials, Transactions of this Society, vol. 31 (1929).

9. Titchmarsh, E. C., *The zeros of certain integral functions*, Proceedings of the London Mathematical Society, vol. 25 (1926), p. 283.

10. Cartwright, M. L., *The zeros of certain integral functions*, Quarterly Journal of Mathematics, vol. 1 (1930), p. 38.

THE UNIVERSITY OF WISCONSIN

TYPES OF SERIES AND TYPES OF SUMMABILITY*

BY C. N. MOORE

1. *Introduction.* In most historical accounts of the theory of divergent series considerable stress is laid on the dictum of Abel and Cauchy with regard to their use and its influence on the study of such series. While not wishing to deny the importance of this influence, particularly in the case of the French school of mathematics, I nevertheless feel that there has been a tendency to exaggerate it. This has doubtless arisen from the fact that the excellent historical discussion given in Borel's *Leçons sur les Séries Divergentes* has generally been accepted as having wider scope than it actually possesses. The account there given of the effect of the Abel-Cauchy dictum is primarily a description of its influence on the French school, although that fact is not explicitly stated. It is to be expected that any interpretation of Borel's remarks as describing the state of affairs in the mathematical world at large would lead to a somewhat warped view of the situation.

Mathematical science, like all other living things, has its own natural laws of growth. I do not believe that the dictum of any two mathematicians, even of the stature of Abel and Cauchy, is sufficient to ban from consideration any particular branch of investigation that possesses intrinsic importance. It is quite true that the attention paid to divergent series steadily diminished during the first eighty years of the nineteenth century, and that it had almost reached the vanishing point during the last decade of that period. This was due, however, to a variety of factors, of which the Abel-Cauchy dictum was only one element. The gradual development of our present day notions of rigor was perhaps the dominant influence. The general standards of rigor set forth by Abel and Cauchy played a very important role in this development, and thus their indirect influence on the study of divergent series was in itself of considerable importance.

* An address presented at the meeting of the Society at Columbia, Missouri, November 29, 1930, by invitation of the program committee.

The extent to which the study of divergent series was continued, both in England and Germany, during the first sixty years of the nineteenth century has been clearly pointed out by Burkhardt in his paper of 1911 in the Mathematische Annalen, *Über den Gebrauch divergenter Reihen in der Zeit von 1750–1860*. Since there was a steady development of notions of rigor during that same period, one may naturally inquire why the rigorous treatment of divergent series did not get an earlier start. I think that the reason is not far to seek. From the point of view of relative frequency and relative simplicity, convergent series bear somewhat the same relationship to series in general that analytic functions do to general functions. Just as it was not feasible for mathematicians to undertake the study of very general types of functions until they had a considerable understanding of the highly special but particularly important class of analytic functions, so it was exceedingly difficult to build a rigorous theory of divergent series before the theory of convergent series had reached some degree of completeness. The attainment of this simpler objective was a task of no mean difficulty, and it absorbed the main energy of students of series during the greater part of the nineteenth century. It is an interesting coincidence that Pringsheim's comprehensive study of convergent series with positive terms was published in the same year (1890) as Cesàro's fundamental paper on the multiplication of series. The latter paper, as we all know, was one of the prime sources of the work of the present century in the field of divergent series.

2. *Origins of the Modern Theory.* Cesàro's paper, published in the Bulletin des Sciences Mathématiques, is just six pages in length and involves only the simplest type of analysis. Nevertheless, if we should take as our criterion of importance for mathematical papers the amount of subsequent literature to which they have led, an examination of the mathematical literature of the first thirty years of the present century would inevitably compel us to rank this paper among the most important contributions of the final quarter of the nineteenth century. At present writing, any such estimate of its importance would appear to be subject to revision upward.

It is also worth while to note, from the standpoint of historical background, that Cesàro clearly understood the scope of the

ideas contained in his paper. Toward the close of it we find the following comments:

"Il est téméraire d'affirmer que les séries non convergentes n'auront jamais d'utilité. Tant que cette assertion restera gratuite, nous serons en droit de rechercher sous quelles conditions on peut soumettre les séries indéterminées aux opérations de l'analyse. Après tout, n'est-ce pas en vertu d'une convention que les séries convergentes, prises sous leur forme indéfinie, interviennent dans les calculs?"

We may state, in passing, that in the second edition of Borel's *Leçons sur les Séries Divergentes*, the remark contained in the last sentence of the above quotation is considered to be of recent origin and is ascribed to Knopp. Thus we see again how even the best expository treatments of a subject may serve to disseminate erroneous notions as to its historical development.

The earlier work of Frobenius and Hölder in connection with the generalization of Abel's theorem on the limiting value of a power series to certain types of divergent series has likewise considerable importance in connection with the modern development of the theory of divergent series. The relationship of this work to Cesàro's and its influence on other important studies in the general field of divergent series has already been adequately discussed in various expository treatments of the subject. The most important feature of this work, from the point of view of the present discussion, is that Hölder's general definition and Cesàro's general definition, which are identical for the simple case discussed by Frobenius, were regarded from the first as being of the same general scope, and their complete equivalence was suspected for some time before it was definitely proved.

3. *Divergent Series and the Problem of Analytic Extension.* It was Borel's attack on the problem of analytic extension from the standpoint of the theory of divergent series which first showed clearly that essentially different types of series would in general require essentially different types of summability. The rather obvious necessary condition for either Cesàro or Hölder summability of order r, namely $\lim_{n\to\infty}[u_n/n^r]=0$, at once shows that these methods will be completely ineffective for summing power series outside of their circle of convergence. Nevertheless, Borel found it useful to take Cesàro's notion of mean

value as a point of departure and to generalize it in such a manner as to obtain an effective method for his purpose. This led to the definition known as the Borel mean. A simple transformation yielded the Borel integral definition, which is of somewhat wider scope. The effectiveness of these methods in connection with the study of power series beyond the domain of convergence can be well illustrated by means of the simple series

$$1 + z + z^2 + z^3 + \cdots .$$

We use the integral definition which is set up as follows. Given any series,

$$u_0(z) + u_1(z) + u_2(z) + \cdots ,$$

we define

$$u(z, t) = u_0 + u_1 t + u_2 \frac{t^2}{2!} + \cdots ,$$

and we form

$$\int_0^\lambda e^{-t} u(z, t) dt = \psi(z, \lambda).$$

If $\lim_{\lambda \to \infty} \psi(z, \lambda)$ exists and is equal to $s(z)$, we say that the series is summable (B) to the value $s(z)$. Applying this to the series $\sum z^n$, we have

$$\psi(z, \lambda) = \int_0^\lambda e^{-t} e^{zt} dt = \int_0^\lambda e^{(z-1)t} dt = \left[\frac{e^{(z-1)\lambda}}{z-1} - \frac{1}{z-1} \right],$$

and it is readily seen that $\psi(z, \lambda)$ approaches $1/(1-z)$ as a limit as λ becomes infinite, provided $R(z) < 1$. Therefore the series $\sum z^n$, which diverges everywhere outside the unit circle, will be summable (B) to the value $1/(1-z)$ for all values of z in the half-plane bounded on the right by the perpendicular to the axis of reals at the point $z = 1$.

Thus we see that the introduction of summability (B) adds very considerably to the region in which the series furnishes the value of the generating function. However, it is apparent that there are still some worlds to conquer, and a more powerful

method is obviously desirable. This was provided by Leroy by means of the following definition. We set

$$F(z, t) = \sum_{n=0}^{n=\infty} \frac{\Gamma(nt + 1)}{\Gamma(n + 1)} u_n(z) \qquad (0 < t < 1), \tag{1}$$

and we define the limit as $t \to 1$ of $F(z, t)$ as the sum of the series $\sum u_n(z)$. Let us apply this method to the series $\sum z^n$. Making use of (1) and the following formula for the gamma function

$$\Gamma(nt + 1) = \int_0^\infty e^{-x} x^{nt} dx,$$

we obtain

$$F(z, t) = \int_0^\infty e^{-x} \sum \frac{(zx^t)^n}{n!} = \int_0^\infty e^{(-x+zx^t)} dx. \tag{2}$$

The path of integration here is along the axis of reals in the x-plane, but this path may be replaced by any other straight line L from the origin to ∞ making an angle less than $\frac{1}{2}\pi$ (in absolute value) with the first path, since the integral along a circular arc joining the two paths approaches zero as the radius becomes infinite.

We can now show by properly choosing the line L, subject to the condition just stated, that the integral in the right-hand member of (2), which converges for $0 < t < 1$ and any fixed value of z, will approach the value $1/(1-z)$ as t approaches 1, for any value of z that does not lie on the axis of reals in the interval $(1 \leqq z < \infty)$, that is for any value of z for which the series $\sum z^n$ does not diverge to $+\infty$. Consider first the behavior of the integral for $t = 1$. If we set $z = \alpha + \beta i$, $x = \rho e^{\phi i}$, it is readily seen that the integral converges to the value $1/(1-z)$, provided z is so chosen that $\alpha \cos \phi - \beta \sin \phi - \cos \phi$ is negative. This condition will be satisfied if the point z lies in a region (T), which is that one of the two half-planes bounded by the line

$$\alpha \cos \phi - \beta \sin \phi - \cos \phi = 0 \tag{3}$$

which contains the origin. We note that the line (3) in the z-plane is the line through the point $z = 1$, parallel to the line symmetric to L with respect to the bisector $\beta = \alpha$.

For any point z in (T) the integral converges uniformly for positive values of $t \leqq 1$; hence the function which it defines, $F(z, t)$, approaches $F(z, 1) = \int_0^\infty e^{z(x-1)}dx = 1/(1-z)$ as t approaches 1. By allowing the line L to take on all admissible positions, the corresponding regions (T) will include any point in the z-plane with the exclusion of points lying in the cut from $z=1$ to $z=\infty$ along the axis of reals.

It would seem at first sight that we should rest content with the result just obtained. We have succeeded in summing the series $\sum z^n$ for all values of z except those for which all the terms are positive and the series diverges to $+\infty$. However, even for such values the series is still the formal expansion of the function $1/(1-z)$ by continued division, by the binomial expansion, or by Taylor's theorem. Why then should it not yield us the sum $1/(1-z)$, if we apply a *suitable* method of summation? This suitable method of summation was furnished by Mittag-Leffler in an article published in volume 42 (1920) of the Acta Mathematica.

We define $E(z)$ by the formula

$$E(z) = \frac{1}{2\pi i}\int^{S} e^{e^{\zeta}} e^{-e^{e^{\zeta}}} \frac{d\zeta}{\zeta - z},$$

where the contour S in the ζ-plane is composed of two lines parallel to the axis of reals and extending to the point at infinity on opposite sides of this axis at a distance between $\pi/2$ and $3\pi/2$, together with a line joining these parallels and cutting the axis of reals at an arbitrary point whose abscissa exceeds the real component of z. The function $E(z)$ can be shown to be an integral transcendental function. We now set

$$G(z) = \frac{E(z+b)}{E(b)},$$

where b is real, and we define $K_n(\omega)$ by means of the expansion

$$G[\omega(x-1)] = \sum_{n=0}^{\infty} K_n(\omega)x^n.$$

It can then be shown that if we put $s_n(z) = 1+z+z^2+\cdots+z^n$,

$$\lim_{\omega \to \infty} \sum_{n=0}^{\infty} s_n(z) K_{n+1}(\omega) = \frac{1}{(1-z)}$$

for every value of $z \neq 1$, and that the left hand side becomes positively infinite, as ω becomes infinite, for the value $z=1$. We can ask no more of a method for summing the formal expansion of $1/(1-z)$.

The various methods of summation which have been illustrated by means of the simple function $1/(1-z)$ and its corresponding Taylor's series apply to much more general cases. Consider any function, analytic in general but having certain singular points, and its corresponding power series development, $u_0+u_1z+u_2z^2+\cdots$. Through each singular point draw a line perpendicular to the line joining the singular point with the origin. These lines form a polygon, known as the Borel polygon, within which the series is summable by Borel's method to its corresponding function. Draw the infinite half-rays which form the prolongation of the lines drawn from the origin to the singular points. The entire plane, with the exception of these rays, forms a region known as the Mittag-Leffler star. For all points within the star the series is summable to the value of the function by Leroy's method. Finally it is summable by Mittag-Leffler's method to the value of the function at all points of the plane except perhaps the singular points themselves.

4. *Summability of Fourier Series and General Orthogonal Developments.* We have seen that Cesàro's methods are ineffective for the study of power series beyond the circle of convergence. They have proved, on the other hand, to be the methods *par excellence* for the study of Fourier series and other developments in orthogonal functions. Their application to problems of this type dates back to Fejér's fundamental paper of 1903, *Untersuchungen über Fouriersche Reihen.* Fejér's theorem concerning the summability $(C1)$ of Fourier series at all points of continuity or discontinuity of the first kind is now regarded as part of the elementary theory of Fourier series. The fact that the Cesàro means form the ideal method for summing Fourier series is better illustrated, however, by the generalizations of Fejér's theorem due to Lebesgue and Hardy. Let us consider the scope of these results. Any function having a Lebesgue integral over a certain interval serves to define for that interval a unique

Fourier development. The same Fourier development, however, will correspond to any other function differing from the first one at any set of points of measure zero. It is clear then that we cannot in general expect the Fourier development of a function having a Lebesgue integral to furnish the value of the function at all points of the interval, but only for a subset differing from the interval by a set of measure zero. Lebesgue showed in 1905 that we could obtain this maximum result by using summability $(C1)$ and Hardy showed in 1913 that we could also obtain it by using summability $(C\rho)$ for any $\rho>0$.

The general state of development of the study of the summability of Fourier series and other expansions in special orthogonal functions that had been reached at the beginning of 1918 was outlined in the writer's Chicago symposium paper, which appeared in volume 25 (1919) of this Bulletin. We shall therefore confine ourselves here to indicating some of the most striking contributions subsequent to that date.

It is well known that the efforts to obtain necessary and sufficient conditions for the convergence of Fourier series have thus far been unsuccessful. The analogous problem for summability by Cesàro means of any particular order is likewise unsolved. To Hardy and Littlewood is due the simple but ingenious idea of formulating a somewhat different problem in the case of Cesàro summability that admits a complete and elegant solution. They considered the situation in which the Fourier series is summable by some Cesàro mean or other, but not necessarily by any particular mean. For this case they obtained the following beautiful result.*

Given a function $f(t)$ that is periodic and has a Lebesgue integral. We set

$$\phi(t) = f(x+t) + f(x-t) - 2A,$$

$$\phi_1(t) = \frac{1}{t}\int_0^t \phi(\alpha)d\alpha, \quad \phi_2(t) = \frac{1}{t}\int_0^t \phi_1(\alpha)d\alpha, \cdots.$$

The necessary and sufficient condition that the Fourier development of $f(t)$ should be summable (C) to the value A at the point

* See Mathematische Zeitschrift, vol. 19 (1923), pp. 67–96.

$t=x$ is that there should exist an integer k for which $\phi_k(t)\to 0$ when $t\to 0$.

The function $\phi_k(t)$ is expressible in the form

$$\frac{k}{t^k}\int_0^t (t-u)\phi(u)du,$$

and in this form the definition is readily extensible to non-integral values of k. The condition $\phi_k(t)\to 0$ with t may be conveniently expressed by saying that $\phi(t)\to O(Ck)$, on account of the analogy between this generalized limit and the Cesàro mean for a series. In addition to their necessary and sufficient condition, Hardy and Littlewood proved the more precise result that if $\phi(t)\to O(Cr)$ when $t\to 0$, r being an integer, the Fourier series is summable $(C, r+1)$ to the sum A for $t=x$, and also that if the Fourier series is summable (C, r) to the sum A for $t=x$, then $\phi(t)\to O\ (C, r+2)$. They showed somewhat later* that if $\phi(t)\to O\ (C\alpha)$ for $0<\alpha<1$, then the series is summable $(C, \alpha+\delta)$ to the sum A at $t=x$ for any positive δ, and also that if the series is summable $(C, -\gamma)$, for $0<\gamma<1$, to the sum A at $t=x$, then $\phi(t)\to O\ (C, 1)$. Quite recently Bosanquet has obtained notable generalizations† of these theorems. He extends the first result quoted above to the case of any non-negative α, and he proves the following generalization of the second result. If the Fourier series is summable $(C\alpha)$ to the sum A at $t=x$, for any $\alpha\geqq -1$, then $\phi(t)\to O\ (C, \alpha+1+\delta)$ for any positive δ.

In the case of developments in general orthogonal functions Cesàro's methods have likewise proved to be of the greatest practical importance. A series of studies by various authors of the scope of these methods has yielded results of gradually increasing generality. The theory has reached finality in the following very general theorem of Menchoff.‡ For any series of orthogonal functions, $\sum a_n\phi_n(x)$, the condition

$$\sum a_n^2(\log\log n)^2<\infty$$

is sufficient for summability $(C, k<0)$ almost everywhere in the

* See Proceedings of the Cambridge Philosophical Society, vol. 23 (1927), pp. 681–684.

† See Proceedings of the London Mathematical Society, (2), vol. 31 (1930), pp. 144–164.

‡ See Fundamenta Mathematicae, vol. 8 (1926), pp. 56–108.

interval of definition of the $\phi_n(x)$. For any positive function $\omega(n)$ satisfying the condition $\lim_{n\to\infty}[\omega(n)/(\log\log n)^2]=0$, there exists a series of orthogonal functions for which $\sum a_n\omega_n<\infty$, whereas the series is not summable by any Cesàro mean.

5. *Summability of Dirichlet's Series.* The methods of Cesàro have also been applied to advantage to certain types of Dirichlet's series. For the general Dirichlet's series, $\sum a_n e^{-\lambda_n s}$, we need the more general methods of summation introduced by Marcel Riesz and known as summation by typical means. Representing the general term of the series by u_n, we set

$$C_\lambda^\kappa(\omega)=\sum_{\lambda_n<\omega}(\omega-\lambda_n)^\kappa u_n,\quad C_l^\kappa(w)=\sum_{l_n<w}(w-l_n)^\kappa u_n,$$
$$(l_n=\log\lambda_n).$$

If the limits $\lim_{\omega\to\infty}[C_\lambda^\kappa(\omega)/\omega^\kappa]$, or $\lim_{w\to\infty}[C_l^\kappa(w)/w^\kappa]$, exist, we say that the series is summable (R,λ,κ) or (R,l,κ), respectively, to the value of the limit in question. The special cases of the Riesz definition in which $\lambda_n=n$ or $l_n=n$ are entirely equivalent to the Cesàro mean of the same order κ. However, even in this special case, it appears that the Riesz means are better adapted to the study of Dirichlet's series. An analogous situation arises in the case of the developments in Bessel's functions, where equivalent Rieszian means can sometimes be used with more facility than the Cesàro mean. Thus we see that even in the case of two equivalent methods of summation, one of the two may be better adapted to the study of a particular type of series.

The domain of effectiveness of the Riesz typical means in summing Dirichlet's series is in general a half-plane bounded on the left by a line parallel to the axis of imaginaries, although it may be the whole plane or a null region. In the case where we have a half-plane of summability, it does not necessarily follow that the finite portion of the boundary line contains a singular point of the function corresponding to the series. In fact this function may be analytic all over the finite plane, and still the series may be non-summable by any typical mean in a certain half-plane. Our success in extending the domain of summability of power series by using more powerful methods suggests that something analogous may be possible here. That this is actually

the case was shown by Riesz in an article that appeared in volume 35 (1912) of the Acta Mathematica.

Riesz took as his point of departure a method of summation used by Mittag-Leffler in connection with power series, which has for such series the same general range of effectiveness as Leroy's method. Adapting this method to the special nature of Dirichlet's series in somewhat the same manner as in the case of his typical means, Riesz sets

$$\phi(\alpha, s) = \sum_{n=0}^{\infty} \frac{1}{\Gamma(\alpha\lambda_n + 1)} a_n e^{-\lambda_n s}$$

and defines $\lim_{\alpha \to 0} \phi(\alpha, s)$ as the generalized sum of the series. He shows that for any Dirichlet's series that has a domain of convergence, this method will sum the series to the value of the function corresponding to the series everywhere in the finite plane, except along half-lines drawn through the singular points in the direction of the negative half of the axis of reals. This region is the analogue of the Mittag-Leffler star for power series.

6. *Conclusion.* In conclusion I wish to point out that I have not attempted to give a comprehensive account of the application of various methods of summation to particular types of series. It would be impossible to give such an account within the reasonable bounds of the ordinary expository article. I have aimed rather to select certain high lights of the theory which seem to me to bear out the main thesis of this paper, namely that the ideal type of summation is necessarily different for different types of series. The entire history of the theory furnishes ample warning against a predisposition in favor of any particular method. In approaching the study of any essentially new type of series, our choice of known methods or our construction of a new method must be governed entirely by the inherent nature of the series in question.

THE UNIVERSITY OF CINCINNATI

FACTORABILITY OF NUMERICAL FUNCTIONS*

BY E. T. BELL

1. *Factorability.* To make this note self-contained we recall a few definitions and theorems from previous papers 1, 2. The new results complete a previous theory in an interesting detail. References are in §4. As some of the conclusions are rather unexpected, we shall state one in full in ordinary notation while giving the definitions; it is restated in §2 and proved in §3 by the symbolic method appropriate to the subject.

If $f(x)$ is single-valued and finite for integer values >0 of x, $f(x)$ is called a numerical function of x. If $k(x)$ is a numerical function of x having the property that

$$k(mn)=k(m)\ k(n)$$

whenever m, n are coprime integers, <0, $k(x)$ is said to be factorable. If $g(x)$ is a numerical function, $g(x)$ is said to be regular or irregular according as $g(1)\neq 0$ or $g(1)=0$.

Let $f(x)$, $g(x)$ be regular numerical functions. Define $h(n)$ by

$$h(n) = \sum f(d)g(t) \tag{1}$$

for all integers n, where the sum refers to all pairs (d, t) of positive divisors of n such that $dt=n$. Then, by the definition, $h(x)$ is a numerical function of x. Among other theorems we shall prove the following.

If $h(x)$ is factorable, then either both of $f(x)$, $g(x)$ are factorable or neither of $f(x)$, $g(x)$ is factorable.

The converse is false, but *if both of $f(x)$, $g(x)$ are factorable, $h(x)$ is factorable.*

As in the papers 1, · · · , 4 cited below, we shall write (1) symbolically, $h=fg$, and refer to f instead of $f(x)$ as a numerical function.

The unit numerical function is the unique η such that $\eta f=f$ for all numerical functions f, and η is defined by $\eta(1)=1$, $\eta(n)=0$ $(n>1)$. If and only if f is regular, there exists[3] a unique

* Presented to the Society, November 29, 1930.

regular numerical function, denoted by f^{-1}, such that $ff^{-1}=\eta$. It is necessary to recall that the symbolic multiplication in $fg\cdots h$, where $f, g, \cdots, h$ are numerical functions, is associative and commutative;[2,4] it is also distributive over the relevant addition,[6] but this is not required here. Further, if r is an integer >0, f^r denotes the product (in the above sense) of r factors f; if s is a rational number, $f^s=g$ is interpreted as $f^a=g^b$, where a, b are coprime integers (positive or negative) and $s=a/b$. If in the last one of a, b is negative, say b, we may write $f^a=g^b$ in the equivalent forms $f^ag^{-b}=\eta$, b positive.

If and only if f is regular and factorable it has a generator[2,5] of the form

$$\Gamma(f)\equiv 1+F_1(p)z+F_2(p)z^2+\cdots,$$

where the series need not terminate, and the F's are single-valued functions of p. For details we must refer to previous papers.[5] If $a, b, \cdots,$ are rational numbers, and

$$\Gamma(f), \Gamma(g), \cdots, \Gamma(h)$$

are the generators of the regular factorable numerical functions $f, g, \cdots, h$, the generator[5] of $fg\cdots h$ is that branch of the formal expansion of

$$(\Gamma(f))^a(\Gamma(g))^b\cdots(\Gamma(h))^c$$

whose term free of p, z is 1. If $f, g, \cdots, h$ are regular, so also is $fg\cdots h$.

2. *Theorems.* The first is implicit in the previous papers, but we state it here for completeness, as it is used in the proofs.

THEOREM 1. *If f is a regular numerical function and r is a rational number, f^r is factorable if and only if f is factorable. If $f, g, \cdots, h$ are regular factorable numerical functions and $a, b, \cdots, c$ are rational numbers, $f^ag^b\cdots h^c$ is factorable.*

LEMMA. *If f, g, h are numerical functions such that f, g are regular and $fg=h$, then either both of f, g are factorable or neither of f, g is factorable.*

This is used in proving the next, which includes the lemma.

THEOREM 2. *If $f, g, \cdots, h, k$ are numerical functions, of which $f, g, \cdots, h$ are regular and k is factorable, such that*

$$f^ag^b\cdots h^c=k,$$

where $a, b, \cdots, c$ are rational numbers, then all or none of $f, g, \cdots, h$ are factorable.

It remains to be seen that the last is a genuine theorem and not a true but vacuous statement.

THEOREM 3. *There exist an infinity of $f, g, \cdots, h$ satisfying either one of the conclusions of Theorem 2.*

3. *Proofs.* Theorem 1 is a mere restatement of the last paragraph of §1. Theorem 2 then follows by applying Theorem 1 and the Lemma to $f^a g^b \cdots h^c$ considered as $f^a(g^b \cdots h^c)$, repeating the argument for $g^b \cdots h^c$, and so on. In the Lemma we have $g = f^{-1}h$, h factorable. Hence by Theorem 1, if f is factorable, so also is $f^{-1}h$, and therefore g is factorable. If f is not factorable and g is factorable, from $f = g^{-1}h$ we get the contradiction that f is factorable. In Theorem 3 the part concerning the first conclusion of Theorem 2 follows from Theorem 1. The second part follows from Theorem 2 and the first part of Theorem 1, applied to $f^c h^a$, $f^{-c}h^b$, where h is factorable and f is not factorable, and a, b, c are rational numbers. For, the product of $f^c h^a$ and $f^{-c}h^b$, neither of which is factorable, is h^{a+b}, which is factorable.

4. *References.*

(1) *Outline of a theory of arithmetical functions in their algebraic aspects*, Journal of the Indian Mathematical Society, vol. 17 (1928). This paper contains references to my earlier papers on the same topic.

(2) *An arithmetical theory of certain numerical functions*, University of Washington Publications in Science, vol. 1, No. 1, 1915.

(3) *A certain inversion in the theory of numbers*, Tôhoku Mathematical Journal, vol. 17 (1920), pp. 221–231.

(4) *Extensions of Dirichlet multiplication and Dedekind inversion*, this Bulletin, vol. 28 (1922), pp. 111–122.

(5) *Euler algebra*, Transactions of this Society, vol. 25 (1923), pp. 135–154.

(6) *Algebraic Arithmetic*. Colloquium Publications of this Society, vol. 7, 1927.

CALIFORNIA INSTITUTE OF TECHNOLOGY

TWO TYPES OF CONNECTED SETS*

BY P. M. SWINGLE

1. *Definitions.* A set W will be said to be *widely connected* if it is connected† and every connected subset is everywhere dense in it.

A connected set W will be said to be *n-point connected*, where n is any given cardinal number, if there does not exist a subset of power n which disconnects W.

A connected point set (or a continuum) W of type T will be said to be a *perfect connected set* (or a *perfect continuum*) *of type* T if every connected subset (or every subcontinuum) of W is of type T. For example a *perfect one-point connected set* is a connected set, every connected subset of which is one-point connected. A *perfect indecomposable continuum* is a continuum every subcontinuum of which is indecomposable.‡

2. *An Example of a Widely Connected Set.* It will now be shown that under certain logical assumptions, including Zermelo's postulate,§ a widely connected set can exist.

THEOREM 1. *Any bounded indecomposable continuum M, lying in a euclidean space, contains a widely connected subset which is everywhere dense in M.*

Let (K) be the set whose elements are the composants¶ of M, these elements being contained but once in (K). It is known

---

* Presented to the Society, October 25, 1930.

† By the notation $W=W_1+W_2$ *separate* is meant that W is the sum of the two non-vacuous, mutually exclusive subsets W_1 and W_2 neither of which contains a limit point of the other. A set W is *connected* if there do not exist subsets W_1 and W_2 such that $W=W_1+W_2$ separate. In this paper a single point will not be considered a connected set. By the notation W' will be meant the set W plus the limit points of W.

‡ See R. L. Wilder, *Characterizations of continuous curves that are perfectly continuous*, Proceedings of the National Academy of Sciences, vol. 15 (1929), pp. 614–621.

§ See Alonzo Church, *Alternatives to Zermelo's assumption*, Transactions of this Society, vol. 29 (1927), pp. 178–208.

¶ For definition and properties see Z. Janiszewski and C. Kuratowski, *Sur les continus indécomposables*, Fundamenta Mathematicae, vol. 1, pp. 217–222.

that the power of (K) is that of the linear continuum.* Let c represent this power. Let (C) be the set containing as elements all continua in our space which separate two points of M. It is known that (C) may be taken so that its power is c.†

Let $R(K)$ be a one-to-one correspondence between the elements of (K) and the real numbers (r), $0 \leqq r \leqq 1$; and let $R(C)$ be a one-to-one correspondence between the elements of (C) and these real numbers (r). Let K_r be that element of (K) which corresponds to r of (r) and let C_r be that element of (C) which corresponds to this same r.

As K_r is dense in M,‡ it contains at least one point of C_r. Let w_r be one of these points. Let $W=(w_r)$, then, be the set which contains as elements one and only one point of every K of (K).

It is known that W is connected§ and dense in M. Let Z be any connected subset of W. Assume that Z is not dense in W. Hence Z' is a proper subcontinuum of M and so is contained in a set K of (K). Thus Z is contained in K also. But this is impossible as $W \times K$ contains only one point. Hence Z must be dense in W and so W is widely connected.

It is known that a plane indecomposable continuum M, containing a widely connected subset W, can be projected upon a sphere and then upon a plane S in such a manner that the projection of W on S is an unbounded widely connected set. The following theorem is of interest although its proof is evident.

THEOREM 2. *If W is an unbounded widely connected set, then W does not contain a bounded connected subset.*¶

* S. Mazurkiewicz, *Sur les continus indécomposables*, Fundamenta Mathematicae, vol. 10, pp. 305–310.

† B. Knaster and C. Kuratowski, *Sur les ensembles connexes*, Fundamenta Mathematicae, vol. 2, p. 253. Indebtedness to this paper for the method of procedure used here is acknowledged as well as to R. L. Wilder, this Bulletin, vol. 33 (1927), pp. 423–427.

‡ Z. Janiszewski and C. Kuratowski, loc. cit., p. 221, Theorem 8.

§ B. Knaster and C. Kuratowski, loc. cit., pp. 233–234, Theorem 37.

¶ See S. Mazurkiewicz, *Sur l'existence d'un ensemble plan connexe ne contenant aucun sous-ensemble connexe, borné*, Fundamenta Mathematicae, vol. 2, pp. 97–103; B. Knaster and C. Kuratowski, loc. cit., p. 244; and G. Poprougénko, *Sur un ensemble connexe plan ne contenant aucune partie connexe bornée*, Fundamenta Mathematicae, vol. 15, pp. 329–336. It is evident that the last two of these examples are not widely connected, as each contains a disconnecting point. Furthermore the first contains such a point as shown by equation 25, p. 100.

3. *Theorems.* The following theorems will hold for any space in which the sets may exist, unless otherwise stated. The truth of the first theorem is evident.

THEOREM 3. *Every widely connected set is a perfect widely connected set.*

LEMMA 1. *If $C-Q$ is disconnected, where Q is finite and C is connected, then C contains a connected subset K and Q contains a point q such that $K-q$ is disconnected and contains $C-Q$.*

THEOREM 4. *Let n be a given positive integer, and W a widely connected set. Then W is a perfect n-point connected set.*

Lemma 1 is seen readily to be true. And Theorem 4 is true. For assume that C is a connected subset of W which is not an n-point connected subset and so contains a finite subset Q of n points which disconnects it. Hence, by Lemma 1, it contains a connected subset K and a point q such that $K-q$ is disconnected. Let $K-q=K_1+K_2$ separate. Then K_1+q is a connected subset of K which is not everywhere dense in it, contrary to Theorem 3. Thus W must be a perfect n-point connected set.

THEOREM 5. *If W is a widely connected set, lying in a euclidean space, then W is punctiform.*

Assume that W is not punctiform and so contains a subcontinuum C. But as C lies in a euclidean space it contains a proper subcontinuum contrary to Theorem 3. Hence W is punctiform.

That an n-point connected set is not necessarily punctiform is seen from the following theorem.

THEOREM 6. *Let n be a given positive integer and let M be a perfect indecomposable continuum.* Then M is an n-point connected set.*

For under the assumption that M is not an n-point connected set there exists, by Lemma 1, a connected subset K and a point q such that $K-q=K_1+K_2$ separate. Hence the indecomposable

* For an example of a perfect indecomposable continuum see B. Knaster, *Un continu dont tout sous-continu est indécomposable*, Fundamenta Mathematicae, vol. 3, pp. 247–286.

continuum K' is the sum of the two proper subcontinua $(K_1+q)'$ and $(K_2+q)'$, which is impossible.

A similar proof establishes the following theorem.*

THEOREM 7. *Let n be a given positive integer. Then a connected subset everywhere dense in an indecomposable continuum is an n-point connected set.*

THEOREM 8. *Let n be a given positive integer and M be a perfect n-point connected set. Then M contains a proper connected subset which disconnects it.*†

For let q_1 and q_2 be any two points of M. Then $(M-q_1)-q_2$ is a connected set such that $M-((M-q_1)-q_2)=q_1+q_2$ separate.

THEOREM 9. *If W is a widely connected set which does not contain a biconnected subset,*‡ *then W is the sum of a countable infinity of mutually exclusive connected subsets.*

As W is not biconnected it is the sum of two mutually exclusive connected subsets, C_0 and U_1, say; likewise U_1 is the sum of two such subsets C_2 and U_2; and $U_i (i=2, 3, \cdots)$ is the sum of two such subsets C_{i+1} and U_{i+1}. Let $C_1=W-(C_2+C_3+\cdots)$. Then as C_1 contains C_0 it is connected and so W is the sum of the mutually exclusive connected subsets $C_1, C_2, \cdots$.

THEOREM 10. *If W is a biconnected subset of an indecomposable continuum M, then W is widely connected if W is dense in M.*

Assume that W is not widely connected, in which case it contains a connected subset C such that C' does not contain W. Suppose that $(W+C)-C'=W-W\times C'=W_1+W_2$ separate. But then the indecomposable continuum M would be the sum of the two proper subcontinua $(W_1+C')'$ and $(W_2+C')'$, as $W'=M$, which is impossible. Thus $W-W\times C'$ is connected. Suppose now that $W-C=Z_1+Z_2$ separate, where Z_1 say con-

* See B. Knaster and C. Kuratowski, *Sur les continus non-bornés*, Fundamenta Mathematicae, vol. 5, p. 37, Corollary, where a related theorem is stated.

† For a study of sets which do not have this property see R. L. Wilder, *On a certain type of connected set which cuts the plane*, Proceedings of the International Congress held at Toronto, vol. 1, 1928, pp. 423–437.

‡ Whether a widely connected set can contain a biconnected subset is a problem of interest in connection with the unsolved problem proposed by C. Kuratowski, Fundamenta Mathematicae, vol. 3, p. 322.

tains the connected set $W-W\times C'$. But since we also know that $M=W'=(W-W\times C')'+C'$, $(W-W\times C')'=M$ and so Z_1 contains Z_2, which is impossible. Therefore $W-C$ is connected and so W is the sum of two mutually exclusive connected subsets, which is a contradiction. Hence W must be widely connected.

OHIO STATE UNIVERSITY

QUADRILATERALS INSCRIBED AND CIRCUMSCRIBED TO A PLANE CUBIC*

BY A. H. DIAMOND

In a paper by M. W. Haskell† the geometrical configurations of triangles inscribing and circumscribing a plane cubic curve have been studied by analytic methods. The purpose of this paper is to examine the properties of quadrilaterals inscribing and circumscribing a plane cubic curve by means of elliptic functions.

The coordinates of a point on the curve will be expressed in terms of Weierstrass' elliptic functions $\wp(u)$ and $\wp'(u)$. It is known that $3n$ points of the cubic are the points of intersection of the cubic with a curve of order n if‡

$$u_1 + u_2 + \cdots + u_{3n} \equiv 0 \mod (\omega_1, \omega_2). \tag{1}$$

The values of the parameters of the vertices of the quadrilaterals are obtained from a consideration of the congruences

$$2u_1 + u_2 \equiv 0, \quad 2u_2 + u_3 \equiv 0, \quad 2u_3 + u_4 \equiv 0, \quad 2u_4 + u_1 \equiv 0,$$

whence

$$15u_1 \equiv 0,$$

or

$$u_1 = \frac{k_1\omega_1 + k_2\omega_2}{15},$$

---

* Presented to the Society, November 29, 1930.

† Haskell, this Bulletin, vol. 25 (1918), p. 194.

‡ Appell and Lacour, *Théorie des Fonctions Elliptiques et Applications*, Chap. 3.

where $k_1, k_2 = 0, 1, \cdots, 14$. There are 225 incongruent values of u which include the 9 values of u corresponding to the points of inflexion. The 216 remaining values give the 216 vertices of the 54 quadrilaterals.

The properties of the quadrilaterals may be derived from the matrix given below in which a quadrilateral of a given column corresponds to the point of inflexion represented in the same column. In this table (i, j) denotes $(i\omega_1 + j\omega_2)/15$.

0, 0	5, 0	10, 0	0, 5	5, 5	10, 5	0, 10	5, 10	10, 10
3, 0	2, 0	1, 0	3, 5	2, 5	1, 5	3, 10	2, 10	1, 10
9, 0	11, 0	13, 0	9, 5	11, 5	13, 5	9, 10	11, 10	13, 10
12, 0	8, 0	4, 0	12, 5	8, 5	4, 5	12, 10	8, 10	4, 10
6, 0	14, 0	7, 0	6, 5	14, 5	7, 5	6, 10	14, 10	7, 10
3, 3	2, 12	1, 6	3, 8	2, 2	1, 11	3, 13	2, 7	1, 1
9, 9	11, 6	13, 3	9, 14	11, 11	13, 8	9, 4	11, 1	13, 13
12, 12	8, 3	4, 9	12, 2	8, 8	4, 14	12, 7	8, 13	4, 4
6, 6	14, 9	7, 12	6, 11	14, 14	7, 2	6, 1	14, 4	7, 7
3, 6	2, 9	1, 12	3, 11	2, 14	1, 2	3, 1	2, 4	1, 7
9, 3	11, 12	13, 6	9, 8	11, 2	13, 11	9, 13	11, 7	13, 1
12, 9	8, 6	4, 3	12, 14	8, 11	4, 8	12, 4	8, 1	4, 13
6, 12	14, 3	7, 9	6, 2	14, 8	7, 14	6, 7	14, 13	7, 4
3, 9	2, 6	1, 3	3, 14	2, 11	1, 8	3, 4	2, 1	1, 13
9, 12	11, 3	13, 9	9, 2	11, 8	13, 14	9, 7	11, 13	13, 4
12, 6	8, 9	4, 12	12, 11	8, 14	4, 2	12, 1	8, 4	4, 7
6, 3	14, 12	7, 6	6, 8	14, 2	7, 11	6, 13	14, 7	7, 1
3, 12	2, 3	1, 9	3, 2	2, 8	1, 14	3, 7	2, 13	1, 4
9, 6	11, 9	13, 12	9, 11	11, 14	13, 2	9, 1	11, 4	13, 7
12, 3	8, 12	4, 6	12, 8	8, 2	4, 11	12, 13	8, 7	4, 1
6, 9	14, 6	7, 3	6, 14	14, 11	7, 8	6, 4	14, 1	7, 13
0, 3	5, 12	10, 6	0, 8	5, 2	10, 11	0, 13	5, 7	10, 1
0, 9	5, 6	10, 3	0, 14	5, 11	10, 8	0, 4	5, 1	10, 13
0, 12	5, 3	10, 9	0, 2	5, 8	10, 14	0, 7	5, 13	10, 4
0, 6	5, 9	10, 12	0, 11	5, 14	10, 2	0, 1	5, 4	10, 7

The diagonals of any quadrilateral of a given column intersect in the point of inflexion of the same column. Three quadrilaterals of a row which correspond to 3 collinear inflexions are perspective in pairs, the centers of perspectivity being the three collinear inflexions. The point of inflexion at which the diagonals of any one of these quadrilaterals meet is the center of perspectivity of the remaining pair. The 12 vertices of 3 such quadrilaterals are the points of intersection of a quartic with the cubic and lie by threes on 12 straight lines, each line containing a vertex of each quadrilateral. If any 3 of these vertices

which are collinear are rejected, the remaining 9 are the points of intersection of a cubic with the given cubic. All of these properties may be obtained by employing the relation (1).

The vertices of 3 quadrilaterals in the same column or 3 quadrilaterals corresponding to 3 collinear inflexions, no two of which are in the same row, lie by threes on 4 straight lines.

Two triangles whose vertices are given by u, $u+\omega_1/3$, $u+2\omega_1/3$, and $-u$, $-u-\omega_1/3$, $-u-2\omega_1/3$, respectively, are seen to be triply in perspective with respect to the 3 collinear inflexions 0, $\omega_1/3$, $2\omega_1/3$. The vertices of the 2 triangles lie on a conic by relation (1). By assigning appropriate values to u it appears that the 12 vertices of 3 quadrilaterals corresponding to the 3 collinear inflexions 0, $\omega_1/3$, $2\omega_1/3$ form 2 pairs of triangles, each pair being triply in perspective with respect to these inflexions. The vertices of the triangles of a pair are opposite vertices of the quadrilaterals. Similar relations exist with respect to any 3 collinear inflexions.

If the cubic has a node the elliptic functions reduce to expressions involving trigonometric functions. If we denote the period by ω, the congruence defining u_1 becomes

$$15u_1 \equiv 0 \mod \omega,$$

whence $u_1 = k\omega/15$, where $k = 0, 1, \cdots, 14$. These 15 incongruent values of u_1 include the parameters of the 3 inflexions. The remaining 12 values give the vertices of 3 quadrilaterals which have the same properties as the first 3 quadrilaterals of the first row in the matrix described above.

If the cubic has a cusp, the elliptic functions reduce to rational functions of u, and there are no quadrilaterals since the congruence does not exist.

THE UNIVERSITY OF CALIFORNIA

ON A TYPE OF ILLUSORY THEOREM CONCERNING HIGHER INDETERMINATE EQUATIONS*

BY E. T. BELL

1. *Introduction.* Let S be a system of equations, some or all of which need not be algebraic, finite or infinite in number, in a finite number μ of independent variables $x_1, \cdots, x_\mu$. If the total number of sets of positive (greater than zero) integral solutions $x_1=x_1', \cdots, x_\mu=x_\mu'$ of S is finite, we shall refer to these sets as all the positive integral solutions of S. If S has an infinity of sets of positive integral solutions, we shall assume that a set of conditions C may be imposed on $x_1, \cdots, x_\mu$ such that S subject to C has only a finite number of sets of positive integral solutions, and we shall refer to these as all the positive integral solutions of S (so that reference to C may be omitted). To indicate the variables when it becomes necessary, we shall write $S \equiv S(x_1, \cdots, x_\mu)$.

Considering the generality of S, one would suspect that it is difficult to state any non-trivial theorem about all the positive integral solutions of S. Nevertheless several theorems concerning these solutions (the immaterial restriction to a finite number of equations being assumed) occur in the literature and have the appearance at first sight of being genuine theorems. The proof of the theorem in §2 reveals its true character; the generalization of this theorem in §3 scarcely needs a proof, while the further generalization in §4, devised for the occasion, betrays the nature of all such theorems. It is to be noticed that the theorems are much worse than trivialities; they direct us to prolix and unnecessary tentative calculations to settle questions whose answers are presupposed from the beginning.

2. *An Illusory Theorem.*† *If in all the positive integral solutions* $x_1, \cdots, x_\mu$ *of* $S(x_1, \cdots, x_\mu)$ *we set* $x_i=d_i\delta_i$, (d_i, δ_i *positive integers*) *in all possible ways, then the number of positive integral solutions* $y_1, \cdots, y_\mu$ *of* $S(y_1^2, \cdots, y_\mu^2)$ *is* $\Sigma\lambda(d_1 d_2 \cdots d_\mu)$,

* Presented to the Society, November 29, 1930.

† Liouville, Journal de Mathématiques, (2), vol. 4 (1859), pp. 271–272; stated without proof; reported in Dickson's *History*, vol. 2, p. 700.

where $\lambda(x) = +1$ *or* -1 *according as the positive integer* x *is a product of an even or an odd number of primes (equal or distinct).*

If the proof had been stated, it would have been seen at once that the theorem is merely an unnecessarily complicated form of the following.

To find the number of positive integral solutions $y_1, \cdots, y_\mu$ *of* $S(y_1^2, \cdots, y_\mu^2)$, *find all the positive integral solutions* $x_i, \cdots, x_\mu$ *of* $S(x_1, \cdots, x_\mu)$, *and actually count those in which each of* $x_1, \cdots, x_\mu$ *is a square.*

If it is not obvious that the method stated in the theorem for finding the required number of solutions is worse than its trivial equivalent, consider a very large x_i, and compare the difficulty of resolving it in all possible ways into the required $d_i\delta_i$ with that of deciding whether or not it is a square by extracting its square root. Both methods presuppose the complete solution of $S(x_1, \cdots, x_\mu)$ in positive integers. The first demands the resolution of all $x_1, \cdots, x_\mu$ in all the solutions into their prime factors, and this is likely to be tentative if only one of the x_i has a large prime factor. In the second, only the extraction of square roots is required, and this is an elementary non-tentative process.

The proof is contained in the mere statement of two properties of $\lambda(x)$: if x, y are positive integers, $\lambda(xy) = \lambda(x)\lambda(y)$, and $\Sigma\lambda(d)$, where the sum extends over all divisors d of x, is equal to 1 or 0 according as x is or is not a square.

Thus the introduction of λ into the problem of finding the number of positive integral solutions of $S(y_1^2, \cdots, y_\mu^2)$ merely gives a uselessly complicated way of recognizing a square. If the theorem states anything that was not already known, it is only the two properties of λ recalled above, both of which are extraneous to the problem.

3. *A Generalization* of the Theorem in* §2. The theorem in §2 can be "generalized" in an endless variety of ways. Thus, consider a positive integral solution $x_1^0, \cdots, x_\mu^0$ of $S(x_1, \cdots, x_\mu)$, and any divisor δ_i^0 of x_i^0. Call the product $\delta_1^0 \cdots \delta_\mu^0$ a divisor product belonging to the solution $x_1^0, \cdots, x_\mu^0$. Then we have the following theorem.

* Gegenbauer, Wiener Sitzungsberichte (Math.), vol. 95, II (1887), pp. 606–609. The above statement is from Dickson's summary, loc. cit., p. 701.

Let $\chi(x)$ be a function for which $\chi(xy)=\chi(x)\,\chi(y)$ for all values of x, y satisfying a definite condition. Let $X(n)=\Sigma\chi(d)$, where d ranges over all divisors of the positive integer n. Then

$$\sum X(x_1^0)\cdots X(x_\mu^0) = \sum\chi(\delta_1^0\cdots\delta_\mu^0),$$

where on the left the summation extends over all those solutions $x_1^0, \cdots, x_\mu^0$ which satisfy the condition mentioned, while on the right the summation extends over all the divisor-products belonging to these solutions.

It is obvious that the conclusion of this theorem is merely a repetition of the definitions of χ, X. The special case $\chi(n)\equiv\lambda(n)$ is the theorem of §2.

4. *Further Generalizations.* Let $\chi_i(x)$ be a function which takes a definite value when x satisfies a definite condition C_i, and define $X_i(n)$ by $X_i(n)=\Sigma\chi_i(d)$, the notation being as in §3. Then, in the same notation,

$$\sum X_1(x_1^0)\cdots X_\mu(x_\mu^0) = \sum\chi_1(\delta_1^0)\cdots\chi_\mu(\delta_\mu^0),$$

where the summation on the right side refers now to all sets $(\delta_1^0, \cdots, \delta_\mu^0)$ formed from the divisor products $\delta_1^0\cdots\delta_\mu^0$. The special case $(\chi_i n)=\chi(n)$, $(i=1, \cdots, \mu)$, χ as in §3, is the theorem in §3. This is only a tautology on the definitions of χ_i, X_i.

Some of the special cases of this triviality are less easily seen through than the general theorem, for example, the following generalization of §2. Let c be a constant positive integer, and n an arbitrary positive integer. Define $\lambda(n, c)$ by $\lambda(1, c)=\lambda(n, 1)=1$, and for $c>1$, $\lambda(n, c)=0$ if any prime divides n to exactly a $(jc+h)$th power, $j\geqq 0$, $1<h<c$; $\lambda(n, c)=(-1)^t$ in the contrary case, where t is the number of distinct primes dividing n to exactly a $(jc+1)$th power, $j\geqq 0, t\geqq 0$.

Write $\Lambda(n, c)=\Sigma\lambda(d, c)$, the sum extending to all divisors d of n. Then we have the following theorem.

If $c_1, \cdots, c_\mu$ are positive integers, the number of positive integral solutions of $S(y_1^{c_1}, \cdots, y_\mu^{c_\mu})$ is $\Sigma\Lambda(x_1, c_1)\cdots\Lambda(x_\mu, c_\mu)$, the sum extending to all positive integrals solution of $S(x, \cdots, x_\mu)$.

For $c_1=\cdots=c_\mu=2$, this is the theorem of §2 and, like it, is illusory.

CALIFORNIA INSTITUTE OF TECHNOLOGY

PARAMETRIC SOLUTIONS OF CERTAIN DIOPHANTINE EQUATIONS*

BY T. A. PIERCE

In this note parametric solutions of certain diophantine equations are given. The method of obtaining the solutions is derived from an equation involving the determinants of certain matrices. It will be recognized that the method is a generalization of the method of Euler and Lagrange which depends on forms which repeat under multiplication. The matrices used in this paper must be such that their forms are retained under matric multiplication and addition. When integer values are assigned to the parameters of our solutions we obtain integer solutions of the particular equation under consideration; however not all integer solutions are necessarily furnished this way.

Let A be a square matrix with elements in a given field; then

$$|A-\lambda I|=f(\lambda)=0$$

is the characteristic equation of A. If the resultant of $f(\lambda)=0$ and another equation $g(\lambda)=0$ be denoted by $R(f, g)$, then it is known† that

$$|g(A)|=R(f, g).$$

If now B is the matrix which has $g(\lambda)=0$ as its characteristic equation, then

$$|f(B)|=R(g, f)$$

and hence we have

$$|g(A)| = (-1)^{mn}|f(B)|, \tag{1}$$

where m and n are the degrees of $f(\lambda)=0$ and $g(\lambda)=0$, respectively. Equation (1) may be identified with diophantine equations of various types as the following examples illustrate.

Let

$$A=\begin{pmatrix} x, & \alpha y, & \beta z \\ \beta z, & x, & \alpha y \\ \alpha y, & \beta z, & x \end{pmatrix}, \qquad B=\begin{pmatrix} u, & iv \\ iv, & u \end{pmatrix},$$

* Presented to the Society, December 30, 1930.

† Frobenius, Journal für Mathematik, vol. 84, p. 11.

where $\alpha=(a^2b)^{1/3}$ and $\beta=(ab^2)^{1/3}$, and where a and b are relatively prime integers and neither has a square factor. Then 1, α, and β constitute the basis for integers of the pure cubic number field defined by $x^3-a^2b=0$, and $|A|$ is the norm of an integer in this field. Again, $|B|$ is the norm of an integer in the field $k(i)$. The characteristic function of A is

$$f(\lambda) = (x-\lambda)^3 + a^2by^3 + ab^2z^3 - 3ab(x-\lambda)yz,$$

and that of B is

$$g(\lambda) = (u-\lambda)^2 + v^2.$$

When calculating $f(B)$, we first expand and then reduce by $g(B)=0$, since B satisfies its characteristic equation. We obtain

$$g(A) = \begin{pmatrix} X, & \alpha Y, & \beta Z \\ \beta Z, & X, & \alpha Y \\ \alpha Y, & \beta Z, & X \end{pmatrix}, \qquad f(B) = \begin{pmatrix} U, & iV \\ iV, & U \end{pmatrix},$$

where

$$\begin{aligned} X &= x^2 + 2abyz - 2ux + u^2 + v^2, \\ Y &= 2xy + bz^2 - 2uy, \\ Z &= 2xz + ay^2 - 2uz, \\ U &= x^3 + a^2by^3 + ab^2z^3 - 3abxyz + (2u-3x)(u^2+v^2) \\ &\quad + (6xu - 3x^2 + 3abyz - 3u^2 + v^2)u, \\ V &= (6xu - 3x^2 + 3abyz - 3u^2 + v^2)v; \end{aligned}$$

and these, by (1), constitute a five-parameter solution of the equation

$$X^3 + a^2bY^3 + ab^2Z^3 - 3abXYZ = U^2 + V^2.$$

If we take

$$A = \begin{pmatrix} x, & -y \\ y, & x \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} u, & 0 \\ 0, & u \end{pmatrix},$$

we find, by calculating $f(B)$ and $g(A)$ and substituting in (1), that the equation

$$X^2 + Y^2 = Z^2 \tag{2}$$

has the two-parameter solution

$$X = (x-u)^2 - y^2 = m^2 - n^2$$
$$Y = 2y(x-u) = 2mn$$
$$Z = (x-u)^2 + y^2 = m^2 + n^2.$$

This is the well known fundamental solution of (2). If

$$A = \begin{pmatrix} x + (-a)^{1/2}y & , & (-b)^{1/2}(u + (-a)^{1/2}v) \\ (-b)^{1/2}(u - (-a)^{1/2}v), & & x - (-a)^{1/2}y \end{pmatrix},$$

$$B = \begin{pmatrix} p + (-c)^{1/2}q & , & (-d)^{1/2}(r + (-c)^{1/2}s) \\ (-d)^{1/2}(r - (-c)^{1/2}s), & & p - (-c)^{1/2}q \end{pmatrix},$$

then (1) becomes

$$(3) \qquad X^2 + aY^2 + bU^2 + abV^2 = P^2 + cQ^2 + dR^2 + cdS^2$$

and this has the eight-parameter solution

$$X = 2x(p-x) + \theta, \qquad P = 2p(x-p) - \theta,$$
$$Y = 2y(p-x), \qquad Q = 2q(x-p),$$
$$U = 2u(p-x), \qquad R = 2r(x-p),$$
$$V = 2v(p-x), \qquad S = 2s(x-p),$$

where

$$\theta = x^2 + ay^2 + bu^2 + abv^2 - (p^2 + cq^2 + dr^2 + cds^2).$$

When arbitrary integral values are assigned to the parameters we obtain a sequence of numbers simultaneously representable by the quaternary forms on either side of (3).

These examples are sufficient to illustrate the nature and extent of our method. It will be noted in particular that the method applies to certain diophantine equations formed by equating the norms of integers in different fields.

THE UNIVERSITY OF NEBRASKA

SUMS OF FOUR OR MORE VALUES OF $\mu x^2+\nu x$ FOR INTEGERS x*

BY GORDON PALL†

1. *Introduction.* My object is to prove the following theorem.

THEOREM 1. *Let $0<\nu<\mu$, $f(x)=\mu x^2+\nu x$. Let T denote the table of all sums of four values of $f(x)$ for integers x arranged in order of magnitude. The largest gap between consecutive entries of T is*

$$(1) \qquad \mu-\nu, \text{ if } \mu \geqq 3\nu/2; \quad 5\nu-3\mu, \text{ if } \mu \leqq 3\nu/2.$$

An immediate corollary is the following result.

THEOREM 2. *Let $0<\nu<\mu$, $s\geqq 4$. The largest gap in the table of all sums of s values of $f(x)$ for integers x is*

$$(2)\ \mu-\nu, \text{ if } s\mu \geqq (s+2)\nu; \quad (s+1)\nu-(s-1)\mu, \text{ if } s\mu \leqq (s+2)\nu.$$

For, if $s\geqq 4$, we need only add $(s-4)f(-1)$ to every entry of T, notice that a gap $\mu-\nu$ actually occurs from $4f(0)$ to $f(-1)+3f(0)$, that no gap greater than $\mu-\nu$ can exceed $5\nu-3\mu-(s-4)(\mu-\nu)$, and that the last number actually occurs, when it is positive, as the gap from $sf(-1)$ to $f(1)+(s-1)f(0)$.

Let us now recall‡ that the only quadratic functions $q(x)$ which are integers $\geqq 0$ for every integer x, and which take the values 0 and 1 for certain integers x, are obtained from the function

$$(3) \qquad \tfrac{1}{2}mx^2+\tfrac{1}{2}(m-2)x,$$

where m is a positive integer, by replacing x by $x-k$ or $k-x$, k an integer. By Theorem 2, the table of all sums of s values of $q(x)$ possesses as its maximum gap the number 1 if $3\leqq m\leqq s+2$, $m-(s+1)$ if $m\geqq s+2$. One corollary is that every integer $\geqq 0$ is a sum of $m-2$ values of (3) for integers x, all but four of which are 0 or 1, at least if $m\geqq 6$; and of four values if $m=3, 4, 5$; (previously proved by Dickson).

* Presented to the Society, November 29, 1930.

† National Research Fellow in Mathematics, California Institute of Technology.

‡ L. E. Dickson, this Bulletin, vol. 33 (1927), p. 714.

Our problem, since we allow x to have all integer values, is very much simpler than Dickson's in this Bulletin (vol. 33 (1927), pp. 713–720; vol. 34 (1928), pp. 63–72 and pp. 205–217).

2. *Proof of Theorem* 1. We require two lemmas.

LEMMA 1.* *The equations*

$$a = x_1^2 + \cdots + x_4^2, \quad b = x_1 + \cdots + x_4, \tag{4}$$

are solvable in integers x_i if and only if

$$a \equiv b \pmod 2, \quad 4a \geqq b^2, \quad 4a - b^2 \text{ not of the form } 4^h(8v+7). \tag{5}$$

LEMMA 2.† *The equation*

$$p = (3x_1^2 + 2x_1) + \cdots + (3x_4^2 + 2x_4)$$

is solvable in integers x_i, for every $p \geqq 0$.

Let B_a denote the largest b, for a given $a \geqq 0$, for which equations (4) are solvable in integers x_i. If $a \not\equiv 0 \pmod 4$, B_a is, by Lemma 1, the largest integer $b \equiv a \pmod 2$ and satisfying (5_2). Hence

$$(B_a + 2)^2 > 4a, \qquad (a \not\equiv 0, \bmod 4). \tag{6}$$

Then all values b for which (4) are solvable are

$$B_a, B_a - 2, B_a - 4, \cdots, -B_a + 2, -B_a. \tag{7}$$

We verify that, if a is odd,

$$B_{a+1} \leqq B_a + 1, \tag{8}$$

for otherwise $B_{a+1} \geqq B_a + 3$, $(B_{a+1} - 1)^2 > 4a$, which contradicts $(B_{a+1})^2 \leqq 4(a+1)$.

CASE I. $\mu \geqq 3\nu$. Then 2ν and $\mu - \nu$ are permitted as gaps in T. In view of (7) we can pass by differences 2ν to $a\mu + B_a\nu$ from any element $a\mu + b\nu$ of T, if a is odd. In view of (8) we can pass from $a\mu + B_a\nu$ to $(a+1)\mu + B_{a+1}\nu$ by an increment $\leqq \mu - \nu$. Trivially,

---

* A modification of a lemma of Cauchy, this is evident from the identity $4a - b^2 = (x_1 + x_2 - x_3 - x_4)^2 + (x_1 - x_2 + x_3 - x_4)^2 + (x_1 - x_2 - x_3 + x_4)^2$.

† This is equivalent to the readily demonstrable fact that every integer $\geqq 4$ and of the form $3p+4$ is a sum of four squares prime to 3.

$$(a+2)\mu - B_{a+2}\nu - \{(a+1)\mu + B_{a+1}\nu\} \leqq \mu - \nu.$$

Hence, by induction from a to $a+2$, we pass throughout the table.

CASE II. $(5/3)\nu < \mu < 3\nu$. Then $\delta = 2\mu - 4\nu$ and $-\delta$ are both $\leqq \mu - \nu$. Hence, if $a \equiv 1 \pmod 4$ and $b - 2 \geqq -B_{a+2}$ we can reach $a\mu + B_a\nu$ from the entry $a\mu + b\nu$ by successive increments $\mu - \nu$, $\mu - \nu$, $-\delta$, over entries of T. Thus we can pass from $\mu - \nu$ to $\mu + \nu$, thence over 2μ, $3\mu - \nu$, $4\mu - 2\nu$, to $5\mu - 3\nu$ and hence to $5\mu + 3\nu$. If $a \equiv 1 \pmod 4$ and $\geqq 5$ we pass from $a\mu + B_a\nu$ to $(a+1)\mu + (B_a - 1)\nu$, $(a+2)\mu + (B_a - 2)\nu$, $(a+4)\mu + (B_a - 6)\nu$, and hence to $(a+4)\mu + B_{a+4}\nu$, completing the induction.

CASE III. Finally we have the interesting case $\mu \leqq 5\nu/3$. Denote by M_p the class of all $\mu a + \nu b$ such that

$$p = 3a + 2b, \quad 4a \geqq b^2, \quad a \equiv b \pmod 2, \quad (5_3). \tag{9}$$

By Lemmas 1 and 2 there exists a solution a, b of (9) for every integer $p \geqq 0$. Table T coincides with the ordered class of all the elements of all classes M_p.

Let A_p, a_p denote, respectively, the largest and the least values of a satisfying (9); and b_p, B_p the corresponding b's. If p is odd the values a of M_p are

$$A_p, A_p - 4, A_p - 8, \cdots, a_p;$$

the corresponding b's being b_p, $b_p + 6$, $b_p + 12, \cdots, B_p$. Then also $b_p \leqq 1$, $B_p \geqq -3$, $a_p \geqq 1$, and

$$4(A_p + 4) < (b_p - 6)^2, \quad 4(a_p - 4) < (B_p + 6)^2, \quad (p \text{ odd}). \tag{10}$$

If p is odd the difference between two adjacent elements of M_p is allowable, since

$$4\mu - 6\nu \leqq \mu - \nu, \text{ and } 6\nu - 4\mu \leqq 5\nu - 3\mu.$$

The increment from an element $a_1\mu + b_1\nu$ of M_q to an element $a_2\mu + b_2\nu$ of M_{q+1} is $\leqq \mu - \nu$ or $5\nu - 3\mu$, respectively, if

(i) $\mu \geqq 1\frac{1}{2}\nu$, $a_2 \geqq a_1 + 1$; or (ii) $\mu \leqq 1\frac{1}{2}\nu$, $a_1 \leqq a_2 + 3$.

Hence the theorem for the present case will follow from the following lemma.

LEMMA 3. *For every even* $p \geqq 2$,

$$A_p \leqq A_{p-1} + 1, \quad a_{p+1} \leqq a_p + 1; \tag{11}$$

$$a_{p-1} \leqq a_p + 3, \qquad A_p \leqq A_{p+1} + 3. \tag{12}$$

The proof of (11_1) is typical. By (9_2) and (9_3) with $p-1$ and p in place of p, $A_p \equiv A_{p-1}+1 \pmod 4$. Hence the contrary of (11_1) would imply $A_p = A_{p-1}+5+4v_1$, and consequently $b_p = b_{p-1}-7-6v_1$, where $v_1 \geqq 0$. Hence, by $4A_p \geqq (b_p)^2$,

$$4(A_{p-1} + 5 + 4v_1) \geqq (b_{p-1} - 7 - 6v_1)^2,$$

contradicting (10_1) with $p-1$ in place of p, since

$$4(1 + 4v_1) \leqq 2(1 + 6v_1)(6 - b_{p-1}) + (6v_1 + 1)^2.$$

THE CALIFORNIA INSTITUTE OF TECHNOLOGY

GROUPS GENERATED BY TWO OPERATORS WHOSE SQUARES ARE INVARIANT

BY G. A. MILLER

It is well known that two operators of order two generate the dihedral group whose order is twice the order of the product of these operators. The groups that can be generated by two operators which have a common square are also well known. The groups considered in the present article are obviously a generalization of these two categories of well known groups. We shall represent their two generators by s and t. From the fact that s^2 and t^2 are invariant operators of the group G generated by s and t it results directly that

$$s^{-1}sts = t^{-1}stt = ts = (st)^{-1}s^2t^2,$$

$$s^{-1}tss = t^{-1}tst = st = (ts)^{-1}s^2t^2.$$

From these equations it follows that the abelian group H generated by s^2, t^2, and st is invariant under G and that its index under G cannot exceed 2.

A necessary and sufficient condition that H be identical with G is that G be abelian and can be generated by the product of two of its operators and the squares of these operators. It is not

difficult to find a characteristic property of the abelian groups which can be thus generated. In a cyclic group we could use the identity for one of these operators and any generator of the cyclic group for the other. Hence we may restrict our attention to the groups of rank 2. If an abelian group of rank 2 and of order 2^m has the generators s and t, then s^2 and t^2 must appear in a subgroup of index 4. As this subgroup contains $(st)^2$ it results that an abelian group of order 2^m cannot be generated by the product of two of its operators and their squares unless it is cyclic. Since an operator of odd order can always be generated by its square, we may state the following theorem.

THEOREM 1. *A necessary and sufficient condition that an abelian group can be generated by the product of two of its operators and their squares is that its Sylow subgroup of order 2^m is cyclic and that the rank of each of its other Sylow subgroups does not exceed 2.*

In view of this theorem we shall confine our attention in what follows to the non-abelian groups which can be generated by s and t, and involve s^2 and t^2 as invariant operators. The abelian subgroup H is therefore of index 2 under G, and the rank of H cannot exceed 3. The problem of determining all the groups which can be generated by two operators whose squares are invariant is therefore included in that of adjoining to the abelian group H an operator which has its square in H and also transforms the operators of H according to an automorphism of order 2. When H is cyclic, then all the groups which can be thus constructed obviously satisfy the condition that they can separately be generated by two operators whose squares are invariant, but this is not necessarily true as regards the groups thus constructed when H is non-cyclic. We shall first determine the total number of such groups when H is cyclic.

It is well known that the group of isomorphisms of a cyclic group of order p^m, p being an odd prime number, is cyclic and hence it involves only one operator of order 2, while the group of isomorphisms of the cyclic group of order 2^m involves three operators of this order when $m>2$, one such operator when $m=2$, and none when $m=1$. If l represents the number of the distinct prime numbers which divide h, the order of H, the number of operators of order 2 in the group of isomorphisms of the cyclic group H is 2^l-1 when h is either odd or divisible by

4 but not by 8; when h is divisible by 8 this number is $2^{l+1}-1$, and when h is twice an odd number it is $2^{l-1}-1$. When h is odd only one group results from such an automorphism of order 2 while there are two and only two such groups whenever h is even.* Hence the following theorem.

THEOREM 2. *The number of non-abelian groups which can be generated by two operators whose squares are invariant and involve a given cyclic subgroup of order h is 2^l-1 when h is odd, $2^{l+1}-2$ when h is divisible by 4 but not by 8, 2^l-2 when h is twice an odd number, and $6\cdot 2^{l-1}-2$ when h is divisible by 8, l being the number of the distinct primes which divide h.*

When H is non-cyclic and of odd order it is known that in any automorphism of order 2 of H a set of independent generators of H can be so selected that every operator of the set corresponds either to itself or to its inverse.† Moreover, if we extend such an H of rank two or three by an operator of order 2 which transforms one and only one of its independent generators into its inverse and is commutative with the remaining generator when H is of rank 2, or with the other two generators when H is of rank 3, there results a non-abelian group which can be generated by two operators whose squares are invariant. To prove this fact let s_1 and s_2 be two conjugates of the given extending operator of order 2, and suppose that they were so selected that their product is the independent generator of H which is transformed into its inverse. For s we may then use the product of s_1 and another independent generator of H while for t we may take s_2 if H is of rank 2 or s_2 into the remaining independent generator of H when H is of rank 3. Hence the following result has been proved.

THEOREM 3. *Every abelian group of odd order whose rank does not exceed 3, when extended by an operator of order 2 which transforms into its inverse one and only one operator of a set of its independent generators, and is commutative with the remaining operator or operators of the set, gives rise to a group which can be generated by two of its operators whose squares are invariant.*

---

* G. A. Miller, Proceedings of the National Academy of Sciences, vol. 14 (1928), p. 819.

† G. A. Miller, Transactions of this Society, vol. 10 (1909), p. 472.

It should be noted that this theorem includes all the non-abelian groups which can be thus generated when the order of H is odd since such a group must contain an operator of order 2 and this operator could not transform into their inverses two of the independent generators of H. To determine the number of the distinct groups which can involve such a given H we may confine our attention to the case when this H is non-cyclic since this number is included in the theorem given above when H is cyclic. It will be convenient to consider first the case when H is of rank 2 and hence at least one of its Sylow subgroups is of this rank. If one of the Sylow subgroups of H contains two equal invariants, then one of its independent generators must appear in one of the independent generators of H while the other appears in the second independent generator of H. The selection of s and t is therefore not affected by such a Sylow subgroup of H. On the other hand, if a Sylow subgroup of H is either cyclic or has two unequal invariants the operators s and t can be selected in two different ways as regards such a subgroup. Hence we have the following result.

THEOREM 4. *The number of the non-abelian groups which can be generated by two of their operators whose squares are invariant and involve a given abelian subgroup of rank* 2, *index* 2, *and of odd order is* 2^l *where* l *represents the number of the Sylow subgroups of this abelian subgroup which are either cyclic or have two unequal invariants.*

When H is of rank 3 at least one of its Sylow subgroups must be of this rank and a set of independent generators of H can be so selected that the remaining operators of G transform one of the operators of the set into its inverse while they are commutative with each of the other two operators of the set, as was noted above. Since the resulting group is affected only by the selection of the independent generator of H which is transformed into its inverse under G, it results that when H involves a Sylow subgroup which has three equal invariants the selection of s and t is not affected thereby. On the other hand, if H involves a Sylow subgroup which has either three distinct invariants, or only two invariants and these are distinct, the selection of s and t can be made in three different ways in regard thereto. Moreover, this selection can always be made in two

different ways as regards a Sylow subgroup which is either cyclic or involves two and only two equal invariants. This establishes the following theorem.

THEOREM 5. *Any given abelian group of odd order and of rank 3 appears as a subgroup of index 2 in* 2^l+3^k *non-abelian groups which can be generated separately by two of their operators whose squares are invariant, where l represents the number of its Sylow subgroups which are either cyclic or have two and only two equal invariants while k represents the number of these subgroups which have either three distinct invariants or have only two invariants but these are distinct.*

It remains to consider the case when H is both non-cyclic and of even order. Before considering this case it may be desirable to direct attention to the construction of two operators of even order which are commutative, independent, and have the property that the order of their product is equal to the order of the product of their squares. To construct two such operators we first write two cycles on distinct sets of letters such that their orders are equal to one-half of the orders of these operators and multiply one of these cycles on the right and the other on the left by a substitution of order 2 which connects each letter of these cycles with a letter which does not appear in either of them. The square of these products will consist of two equal cycles, one of which is the original cycle. The product of these products will be of the same order as the product of their squares and will be independent of the latter product.

Let H be any abelian group which has at least one even invariant and suppose that the corresponding independent generator of H is associated with st, while the other two are associated with s^2 and t^2, respectively. We may then find two operators of order 2 whose product is the independent generator associated with st and multiply one of these into an operator which is commutative with it and independent thereof such that its square is s^2. This product may be used for s. In a similar way the operator which may be used for t is selected as regards the second of the two given operators of order 2. When H is of rank 2 it may be assumed that $t^2=1$. Hence we have the following fact.

THEOREM 6. *Every abelian group whose rank does not exceed 3 and whose order exceeds 2 is a subgroup of index 2 of a non-abelian group which can be generated by two operators whose squares are invariant under the group, and every group which can be thus generated involves such an abelian subgroup.*

From this theorem it results that the category of groups which are characterized by the fact that each of them can be generated by two of its own operators whose squares are invariant under the group is very large. All such groups are solvable and have a simple structure. For the sake of illustration it may be desirable to note here the smallest of these groups in which H has three invariants. Hence the order of G could not be less than 16 and if G is of this order H must be of type (1,1,1). It is well known that the group of isomorphisms of this H is the simple group of order 168 and that all of its operators of order 2 are conjugate. Hence we have to take for s an operator of order 4 which transforms H according to the said operator of order 2. Since t is also of order 4 and its square is different from s^2 and the commutator of G is s^2t^2, the group G is completely determined. Hence there is one and only one non-abelian group of order 16 which can be generated by two of its operators whose squares are invariant and involves an abelian subgroup of rank 3. It may be noted that this is also the only group of order 16 which is decomposable into two non-invariant cyclic subgroups. In fact, there is one and only one group of order p^4 which is thus decomposable, p being any prime number.*

THE UNIVERSITY OF ILLINOIS

* G. A. Miller, Proceedings of the National Academy of Sciences, vol. 16 (1930), p. 527.

ON A PROBLEM OF N. ARONSZAJN AND AN AXIOM OF R. L. MOORE*

BY LEO ZIPPIN†

In an interesting paper N. Aronszajn‡ raises the question whether every space belonging to a class of topologic spaces which he introduces§ is necessarily an absolute G_δ-set; that is, one which in every metric space M containing it may be expressed as the product of a countable set of sets open in M. In this, the first part of the paper gives an affirmative answer. The class of spaces $\mathfrak{R}$ is given, in effect, by the following axiom: a topologic space R belongs to the class $\mathfrak{R}$ if and only if there exists an infinite sequence of classes $\Pi_R^n(R)$ for $n=1, 2, \cdots$, such that

(1) for every natural number n, R is generated by $\Pi_R^n(R)$;

(2) with every sequence of sets $\{\overline{U}_n\}$, where $U_n \supset U_{n+1}$, there is associated a point p of space such that $p=\Pi_{i=1}^{\infty}\overline{U}_i$ and this sequence converges to p.

The necessary definitions are these: a class $\Pi_R(R)$ is any class of point sets of R which are open in R and whose sum covers R. Such a class generates R if, further, given any point p of R and any open set U containing p, then there is an element V of $\Pi_R(R)$ which contains p and is contained in U. A set of sets converges to a point if every open set containing the point contains all but a finite number of the given sets.

We suppose then that R is a space as defined above and is a subset of a metrizable space M. Then whatever metric be associated with M (preserving its topology) this is induced on R as a subspace (preserving *its* topology) and we may speak of the diameters of the elements of any of the generating classes. Let U be an element of one of these classes whose diameter is finite

---

* Presented to the Society, December 31, 1930.

† National Research Fellow.

‡ N. Aronszajn, *Über die Bogenverknüpfung in topologischen Räumen*, Fundamenta Mathematicae, vol. 15 (1930), pp. 228–241.Compare the first paragraph of page 229.

§ Part of this paper is concerned with the relation of these spaces to spaces defined by R. L. Moore considerably earlier.

(the restriction is convenient) and suppose its diameter, ϵ, not zero. For each point k of U there is in M a spherical neighborhood of k of diameter at most $\epsilon/3$ which contains no point of R not in U, since U is open in R. The sum of all such spherical neighborhoods for every point k of U gives a set U^* which is open in M, of diameter less than 2ϵ, and such that $U=R\cdot U^*$. If ϵ is zero, and U reduces to a point k of R, we shall suppose U indefinitely repeated and associate with it an infinite set of spherical neighborhoods U^* of diameters converging to zero such that $R\cdot U^*=k$. We shall call U^* the extended element of that class of which U is an element. If U, above, is vacuous, it is not relevant to the definition of limit points in R, and is ignored.

Let $F_{n,m}$ be the set of all points of M *which belong to no extended element of the nth generating class which is of diameter less than* $1/m$. This is seen to be closed in M since the extended elements are open there. Then

$$F = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} F_{n,m}$$

is the sum of a countable set of closed sets, and $G=M-F$ is a G_δ-subset relative to M. If, now, x is any point of R and m and n are any preassigned natural numbers, let U_x^* be a spherical neighborhood of x in M of radius less than $1/(4m)$. Since $U_x=R\cdot U_x^*$ is open in R there is an element U of the nth generating class containing x and contained in U_x, and therefore of diameter less than $1/(2m)$. Consequently this has a corresponding extended element U^* of diameter less than $1/m$, and U^* contains x which cannot, therefore, belong to $F_{n,m}$. Since this is true for every point of R, and every n and m, R must belong to G. Suppose, on the other hand, that y is a point of G; then it belongs to no $F_{n,m}$ and is contained in an extended element of any preassigned generating class which is, moreover, of diameter less than any preassigned positive number, not zero. Then y belongs to an extended element U_1^* which is of diameter less than 1. There is a spherical neighborhood U_y^* of y in M such that $U_y^* \subset U_1^*$. Let $1/m_1$ be less than half the radius of U_y^*; y belongs to an extended element U_2^* of the second generating class which is of diameter less than $1/m_1$, and $U_2^* \supset U_1^*$. Continuing inductively, we see that there exists a sequence of extended elements, $U_1^*, U_2^*, \cdots, U_n^*, \cdots$, each containing y, of diameters converging to zero,

each a subset of its predecessor, and each an extended element of the generating class of corresponding number. But $U_n = R \cdot U_n^* \supset R \cdot U_{n+1}^* = U_{n+1}$. Then, by the second condition of the axiom, the set of elements U_n must converge to a point y' of R, and this must be identical with y. Then every point of G is a point of R, and it is shown that R is a G_δ-subset of M.

We shall prove that in metrizable topologic spaces the axiom of N. Aronszajn is equivalent to the following earlier axiom of R. L. Moore.† For every natural number n, $G_1, G_2, \cdots, G_n, \cdots$ is a collection of regions such that (1) for every n, G_n covers space, (2) if R is a region and p and q are points of R, there is an integer d such that if $n \geqq d$ and R_n is a region of G_n containing p, then $\overline{R}_n \subset R - q$, (3) if $R_1, R_2, \cdots$ is a set of regions such that for each n, $R_1, R_2, \cdots, R_n$ have a common point, then there is a point p such that $p \subset \Pi_{i=1}^{\infty} \overline{R}_i$.

Let R be a space satisfying the axiom of Aronszajn and, moreover, metric. We shall define a sequence of generating classes, $G_1, G_2, \cdots, G_n, \cdots$, composed of elements of the corresponding index. It is to be borne in mind that since every spherical neighborhood U_p of a point p of R is open in R, there exists in whatever generating class we wish an element containing p and contained in U_p. Let G_1 be the set of all elements of $\Pi_R^1(R)$ of diameter less than 1. For each point p of R there is some element U_p^1 of G_1 which contains it. There is a lower limit ϵ_p, distinct from zero, to the distances from the point p to the points of $R - U_p^1$. With p we associate all those elements of $\Pi_R^2(R)$ which contain p and whose diameter is less than $\epsilon_p/4$. The set of all such elements, corresponding to each point of R for some element of G_1 arbitrarily associated with it, as above, is the set G_2. Analogously, for every n we form G_n from among the elements of $\Pi_R^n(R)$, with respect to the elements of G_{n-1}. That each G_n is a generating class for R, and that in their totality they form a system equivalent to the original (define the same topologic space and have the properties of the original system of elements) is clear. Every element of G_n, for every n, we shall

† With this axiom and the consequences Moore deduced from it, Aronszajn appears not to be acquainted. Abstract, this Bulletin, vol. 33 (1927), p. 141. A slightly different version of this axiom was presented by R. L. Moore at the Boulder Colloquium Lectures in August of 1929, but the two are readily shown to be equivalent.

call a region (which is an undefined concept in Moore's axiom) and these only shall be regions.

The first condition of Moore's axiom is obviously satisfied by this sequence of regions, and the second by virtue of the fact that a region of G_n is of diameter at most $1/4^{n-1}$. Suppose now that we have a set of regions $R_1, R_2, \cdots$, satisfying the requirements of the third condition of Moore's axiom. Since R_2 is a region of G_2, by the construction of G_2, there exists a point p of R_2 and a region U_p^1 of G_1 such that*

$$\delta(R_2) < \epsilon_p/4 = \rho(p, R - U_p^1),$$

and $\delta(U_p^1) \geqq \epsilon_p$. It is seen that $\rho(R_2, R - U_p^1) \geqq 3\epsilon_p/4$. Let K_2 be any region of G_2 which contains R_3 and let us suppose that K_2 is not contained in U_p^1. Since $K_2 \supset R_3 \supset R_3 \cdot R_2 \neq 0$, and also $K_2 \cdot (R - U_p^1) \neq 0$, it follows that $\delta(K_2) \geqq 3\epsilon_p/4$. Then there is a point q of K_2 and a region U_q^1 of G_1 containing K_2, such that $\delta(K_2) < \epsilon_q/4$, and it follows that $\delta(U_q^1) > 3\epsilon_q$. Since the set of diameters of regions of G_1 is limited above, this process of selecting successive regions of G_1 must terminate with a region P_1 of G_1 with the following properties: there is an integer n_1 such that if $m \geqq n_1$, $P_1 \supset R_m$; and further if Q_2 is any region of G_2 containing R_m, then $Q_2 \subset P_1$. It is possible, inductively, to determine a set of regions, $P_1, P_2, \cdots, P_n, \cdots$, and a set of integers $n_1, n_2, \cdots$, such that (1) P_n is a region of G_n, (2) $P_n \supset P_{n+1}$, (3) if $m \geqq n_k$, then $R_m \subset P_k$. From the first two of these properties it follows (by the second condition of Aronszajn) that the regions (elements) P_n converge to a point p of R. Then for every region $P \supset p$, there is an integer k such that if $m \geqq k$, $P \supset P_m$. By the third property above, $P \supset R_j$ where $j \geqq n_m$. But since $R_i \cdot R_j \neq 0$, where $i = 1, 2, \cdots, j$, it follows that P contains a point of R_i, $i = 1, 2, 3, \cdots$, and $p \subset \overline{R}_i$. Then we have shown that R satisfies the axiom of Moore.

We may define a class of topologic spaces, in a sense the most general, in which an Aronszajn space is necessarily a Moore space, by adjoining to conditions (1) and (2) of Moore's axiom a third to the effect that to a set of regions as described in Moore's axiom may be associated a set $P_1, P_2, \cdots$, with properties as above. We have seen that all metric spaces are in-

* Read: the diameter of R_2 less than lower limit of distances of p to points of $R - U_p^1$.

cluded among these and it may be shown that this class comprises all Moore spaces, and these include examples of non-metric spaces. And it is true, further, that a space of Moore is a space of Aronszajn always. But the author's proofs of these statements he finds to be implicit in various arguments which R. L. Moore has already used (in the pages of his work to appear, in the near future, in the regular Colloquium Series) and in consequence no proof will be given.

The following example of a space satisfying the axiom of Aronszajn (given by him in his paper, page 239) will be used to show that his axiom is more general than Moore's. Let T be the set of all ordinal numbers of the first and second class, where for $n=1, 2, 3, \cdots$, the nth neighborhood of a, $U_n(a)=a$ if a is an isolated ordinal, while if a is a limit-ordinal $U_n(a)$ is a set of numbers z such that $b_n<z\leqq a$, wher $\lim_{n\to\infty}(b_n)=a$. Now, however region be defined or this space in accordance with Moore's axiom, for every limit number p there must occur in one of the covering sets a region R_p which contains none of the successors of p, and R_p must contain infinitely many of the predecessors of p; in particular, one other number q. Then we may associate with p a number n_p such that every region of the n_pth covering set which contains p is contained in R_p, and contains therefore no successor to p. But the set of limit-numbers being uncountable, to uncountably many (in particular to a countable, infinite subset) must correspond the same integer N. Then there is an infinite set (x_n) of these numbers, with a limit number x which succeeds all of them. A region of the $(N+1)$th covering set containing x must contain at least one point of (x_n) contradicting the choice of N. From the self-compactness of the space T, the third condition of Moore's axiom is necessarily fulfilled; but the question whether the fundamental difference between the axioms is constituted by the "uniformity of regularity" which inheres in Moore's second condition (contradicted in the space T) is not settled.

The relation between the axioms is the more challenging since they were fashioned with the same end in view and Moore and Aronszajn drew from them, for their respective spaces, the same impressive consequence.

THE UNIVERSITY OF TEXAS

STATISTICS OF A SET OF CLOSED INTERVALS

BY F. H. MURRAY

1. *Introduction.* In several types of engineering problems it is necessary to consider a group of infinite sequences of time intervals

$$t_{2i+1}^{(k)} \leqq t \leqq t_{2i+2}^{(k)}, \qquad (k = 1, 2, \cdots, m),$$

concerning which little is known except the average number of intervals per unit time, and the frequency distribution of the lengths, for each sequence. Examples are telephone calls being handled in an exchange, and many problems associated with the blowing of fuses.* It is proposed here to develop certain general formulas for a single sequence, and applications to a group of sequences in which the intervals of two distinct sequences are mutually independent.

2. *Non-Overlapping Intervals.* Consider any set of non-overlapping intervals

$$t' \leqq t \leqq t'',$$

and let $2Tn_T(l)$ be the number of intervals of the set each of length $\bar{l}$ greater than or equal to l, whose end points are interior to the interval $-T<t<T$. If from each of these intervals an interval of length l is subtracted on the right, $(t_1'' = t'' - l)$, the sum of the lengths of the parts remaining is a function of l, $2Tp_T(l)$:

$$(1) \qquad p_T(l) = \frac{1}{2T} \sum_{-T<t',t''<T} [\bar{l}(\geqq l) - l] = -\int_l^\infty (s - l)dn_T(s).$$

The functions $n_T(l)$, $p_T(l)$ are non-increasing; and since

$$2Tln_T(l) \leqq \sum_{-T<t',t''<T} (\bar{l} \geqq l) \leqq 2T,$$

it results that $n_T(l) \leqq 1/l$.

Consider any sequence $T_1 < T_2 < \cdots < T_n < \cdots$, $T_n \to \infty$. Since $n_{T_k}(l)$ is bounded for each $l>0$, if l has a rational value a sub-sequence $(T_{k'})$ exists such that

* See T. C. Fry, *Probability and its Engineering Uses*, Chapter 10.

$$\lim_{T_{k'}\to\infty} n_{T_{k'}}(l) = n(l).$$

Then on the set of positive rational numbers not including zero the Cantor diagonal method defines a sequence $(T_{k''})$ and a function $n(l)$, l rational, for which

$$\lim_{T_{k''}\to\infty} n_{T_{k''}}(l) = n'(l).$$

Since $n_T(l)$ is non-increasing, the sequence $(T_{k''})$ also defines a limit function at the points of continuity of $n(l)$.* The points of discontinuity of $n(l)$ form a denumerable set; for since $n_T(l) \leqq 1/l$, the points of discontinuity for which $l \geqq 2^{-m}$ form a denumerable set; and the set of discontinuities in each interval $2^{-m-1} \leqq l < 2^{-m}$ is a denumerable set. The Cantor diagonal method can therefore be employed to define a subsequence $n_{T_k}(l)$ which converges at the points of discontinuity also; and

$$\lim_{T_{k'''}\to\infty} n_{T_{k'''}}(l) = n(l), \quad l > 0,$$

$$ln(l) \leqq 1.$$

Since $n_T(l)$ is non-increasing and bounded for $l \geqq l_1 > 0$, the theorem of Helly can be applied to (1), if $n_T(l) = 0$, $l > L$:

$$\lim_{T_{k'''}\to\infty} p_{T_{k'''}}(l) = p(l) = -\int_l^L (s - l)dn(s). \tag{2}$$

Integrating by parts, we find

$$p(l) = \int_l^L n(s)ds, \tag{3}$$

whence we see that $p(l)$ is continuous.

If no constant L exists for which $n_T(l) = 0$, $l > L$, an integration by parts of (1) gives

$$p_{T_{k'''}}(l) = \int_l^\infty n_{T_{k'''}}(s)ds;$$

and if $l < l'$,

$$-\Delta_l^{l'} p_{T_{k'''}} = p_{T_{k'''}}(l) - p_{T_{k'''}}(l') = \int_l^{l'} n_{T_{k'''}}(s)ds.$$

* For a complete discussion of a similar question, see A. Wintner, *Spektral Theorie der Unendliche Matrizen*, pp. 80–83; Leipzig, 1929.

The theorem of Lebesgue on the passage to the limit under the integral sign gives

$$-\Delta_l^{l'} p = \int_l^{l'} n(s)ds. \tag{4}$$

Since for every T,

$$0 \leqq p_T(l) - p_T(l') \leqq 1,$$

we have also $0 \leqq -\Delta_l^{l'} p \leqq 1$, and as this is a non-decreasing function of l' the integral

$$\int_l^{\infty} n(s)ds \tag{5}$$

converges. This may be taken as the definition of $p(l)$. A sufficient condition for convergence to the limit is the assumption that for any $\epsilon > 0$, an L exists such that

$$p_T(l) < \epsilon, \quad l > L. \tag{a}$$

A related function is $f(l)$, defined by

$$f_T(l) = \frac{1}{2T} \sum_{-T<l',l''<T} \bar{l}(\geqq l) = -\int_l^{\infty} s dn_T(s); \tag{6}$$

from (1),

$$p_T(l) = f_T(l) - l n_T(l).$$

Hence if $p_{T_k'''}$ and $n_{T_k'''}$ approach limits, $f_{T_k'''}$ approaches a limit function $f(l)$, and

$$p(l) = f(l) - l n(l). \tag{7}$$

If $p(l)$ is defined as above when hypothesis (a) is not satisfied, $f(l)$ may also be defined by this equation. In either case, $n(l+0)$ and $n(l-0)$ exist, and from (3)

$$n(l+0) = -\frac{d^+ p(l)}{dl}, \quad n(l-0) = -\frac{d^- p(l)}{dl}, \quad l > 0.$$

Since $p(l) \leqq 1$, the integral (5) converges if the lower limit is zero; $f(l)$ is also bounded and non-increasing; hence $f(0)$ exists. If $n(0)$ exists, $p(0) = f(0)$; in any case

$$p(0) = f(0) - \lim_{l \to 0} l n(l).$$

3. *Several Sets of Intervals.* In what follows it will be assumed that the set of intervals is such that the functions $f(l)$, $n(l)$, $p(l)$ are uniquely defined as T becomes infinite, while $n(0)=n$ is finite. In this case n may be interpreted as the average or expected number of intervals per unit time; the function $g(l)=n(l)/n$ is the cumulative frequency function for the lengths of the intervals. The function $f(l)$ is the probability that a point selected at random falls on an interval of length at least l, and $p(l)$ is the probability that a random point t^0, with all points of the interval (t^0, t^0+l), lies on some interval of the sequence. The number n and the function $g(l)$ determine the remaining functions, and in many cases are the quantities most easily measured.

Suppose two sets present, having end-points (α', β'), (α'', β'') respectively, and assume that the intervals of each set are independent of the other. The set of intervals common to both sets may be divided into two types; in the first, the initial point α' of the first set lies on an interval (α'', β''), and in the second α'' lies on an interval (α', β'). The average number of intervals of length at least Δ, of type (1), is equal to $n'(\Delta)p''(\Delta)$, or the number $n'(\Delta)$ of intervals (α', β') times the probability that α' lies on an interval (α'', β'') together with the interval $(\alpha', \alpha'+\Delta)$. The number of intervals of type (2) is $n''(\Delta)\,p'(\Delta)$; hence

$$N_{12}(\Delta) = n'(\Delta)p''(\Delta) + n''(\Delta)p'(\Delta)$$

is the expected number per unit time. Also, from the general formulas,

$$\begin{aligned} P_{12}(\Delta) &= \int_{\Delta}^{\infty} N_{12}(s)ds \\ &= -\int_{\Delta}^{\infty} [p''(s)dp'(s) + p'(s)dp''(s)] \\ &= -\Big[p'(s)p''(s)\Big]_{\Delta}^{\infty} = p'(\Delta)p''(\Delta), \end{aligned}$$

if the variation of each of the functions (p', p'') over the set of discontinuities of the other is zero.*

* See Hobson, *Functions of a Real Variable*, vol. 1, p. 542, 1927. In many applications this condition can be satisfied by redefining, if necessary, the functions p', p'' in such a way that the original functions are obtained by a passage to the limit.

If there are m sets of intervals, the function $P_{12\cdots m}(\Delta)$ is proved by induction to be equal to $p'(\Delta)p''(\Delta)\cdots p^{(m)}(\Delta)$. The number of such intervals per unit time, of length $\geqq\Delta$, is found by differentiation, and the function $F_{12\cdots m}(\Delta)$ is also determined by the general formulas. The function $N_{12\cdots m}(\Delta)$ can have discontinuities at the values $\Delta = l_i$ which are discontinuities for any of the original sequences, and only at these points, since

$$\frac{d}{d\Delta}P_{12\cdots m}(\Delta)$$

is uniquely determined everywhere else.

4. *Intervals of Constant Length.* If the intervals are of constant length h,

$$\begin{aligned} n(l) &= n(0) = n, \quad l \leqq h, \\ &= 0, \qquad\quad l > h, \\ p(l) &= n(h - l), \quad l \leqq h. \end{aligned}$$

If m sequences are present, with the same n and h,

$$P_m(l) = p(l)^m = n^m(h - l)^m, \quad l \leqq h,$$

$$N_m(l) = mn^m(h - l)^{m-1}.$$

The function

$$\frac{N_m(l)}{N_m(0)} = \left(1 - \frac{l}{h}\right)^{m-1}$$

represents the cumulative frequency distribution for the common lengths, of which the average length is found to be h/m.

If the function

$$g(l) = \frac{n(l)}{n}$$

varies exponentially, we have

$$n(l) = ne^{-l/h},$$

$$p(l) = \int_l^\infty n(s)ds = nhe^{-l/h};$$

and if m similar sequences are present, we have

$$P_m(l) = (nh)^m e^{-ml/h},$$

$$N_m(l) = \frac{m}{h}(nh)^m e^{-ml/h},$$

$$G_m = \frac{N_m(l)}{N_m(0)} = e^{-ml/h}.$$

For a single sequence, the average holding time is

$$-\int_0^\infty l\,de^{-l/h} = h.$$

The average length of an interval common to m sequences is

$$-\int_0^\infty l\,dG_m = \frac{h}{m}.$$

In all cases, the average length of an interval common to m sets is

$$\bar{l} = -\int_0^\infty l\,dG_m(l) = -\frac{1}{N_m(0)}\int_0^\infty l\,dN_m(l)$$

$$= \frac{F_m(0)}{N_m(0)} = -\frac{P_m(0)}{P_m'(0)}.$$

This becomes

$$\frac{1}{\bar{l}_m'} = -\left.\frac{\frac{d}{dl}(p_1p_2\cdots p_m)}{p_1p_2\cdots p_m}\right|_{l=0}$$

$$= -\sum_{k=1}^{m}\frac{p_k'(0)}{p_k(0)} = \sum_{k=1}^{m}\frac{1}{\bar{l}_k}.$$

In many applications the constant $p_k(0) = p_k$ can be interpreted as a load due to an individual source; the equation $\bar{l} = -p(0)/p'(0)$ becomes $\bar{l} = p/n$, and in the form $p = n\bar{l}$ is self-evident.

AMERICAN TELEGRAPH AND TELEPHONE COMPANY

EXTENSION OF A THEOREM OF MAZURKIEWICZ*

BY R. L. WILDER

S. Mazurkiewicz,† in answer to a question proposed by B. Knaster,‡ has shown that if A is a closed point set in E_n (euclidean space of n dimensions) which is homeomorphic with a subset of E_{n-1}, then all points of A are accessible from the complementary set, $E_n - A$. The question naturally arises, then, as to whether the points of A are *regularly* § accessible from $E_n - A$. It will be shown in the present paper that this is indeed the case.

We shall precede our proof by two theorems which, we believe, are themselves of fundamental importance. Following Mazurkiewicz' notation, we shall denote by $S_n(p, \rho)$ a spherical neighborhood of a point p of E_n with radius ρ; by $\phi(A)$, the subset of E_{n-1} that is homeomorphic with A; and if X is any subset of A, by $\phi(X)$ we denote that subset of $\phi(A)$ that corresponds to X under the homeomorphism between A and $\phi(A)$. Also, following the usual custom, if M is a point set, by $\overline{M}$ we shall denote the set M together with all of its limit points.

Evidently the proof given by Mazurkiewicz for his *Lemme* establishes the following more general lemma.

LEMMA 1. *Let P be a point of A, D a domain¶ containing P, and G a component of $D - A \cdot D$ such that $\overline{G} \supset P$. Then, if D_1 is a bounded domain such that $D_1 \subset D$ and $D_1 \supset P$, there is a component G_1 of $G \cdot D_1$ such that $P \subset \overline{G}_1$.*

---

* Presented to the Society, August 30, 1929.

† *Sur un problème de M. Knaster*, Fundamenta Mathematicae, vol. 13 (1929), pp. 146–150.

‡ See Fundamenta Mathematicae, vol. 8 (1926), Problem 43, p. 376.

§ A point P of a point set M is said to be *regularly* accessible from a point set R of which P is a limit point provided that for every $\epsilon > 0$ there exists a positive number δ such that if Q is a point of R whose distance from P is less than δ, then there is an arc from P to Q whose diameter is less than ϵ and which lies, except for P, wholly in R. See G. T. Whyburn, this Bulletin, vol. 34 (1928), p. 509.

¶ By *domain* we mean a connected open subset of the space under consideration. The domain D may, of course, be E_n, in which case the component G of this lemma will necessarily exist, due to the invariance of dimensionality under analysis situs transformations.

THEOREM 1. *If D is a bounded domain of E_n, and a component C of $A \cdot D$ separates* D, then the set $\phi(C)$ is a domain of E_{n-1} whose boundary is $\phi(\overline{C}-C)$.*

PROOF. Since, due to similarities between the proof and that given by Mazurkiewicz for his *Lemme*, we can conserve space by referring to the latter, we shall endeavor to retain most of his notation in this connection.

We can assume that $A=\overline{C}$; then the set $\phi(A)$ does not fill up E_{n-1}. Hence E_n-A is connected,† and if a is a point of C there exists, by the above Lemma, a component G of the set

$$(E_n - C)\cdot D \tag{1}$$

such that $\overline{G} \supset a$. Let $B=A\cdot F=\overline{C}-C$,‡ where F is the boundary of D, and let H be the component of $E_{n-1}-\phi(B)$ that contains $\phi(a)$. We shall show that $H=\phi(C)$.

Since C separates D, there are two points, c and c_1, in (1), which do not lie in the same component of (1); we may suppose that $G \supset c$. Then $\overline{G}-G$ is a cut of E_n between c and c_1, and accordingly contains an *irreducible cut*, L, of E_n, between c and c_1. Let

$$L = L_1 + L_2, \tag{2}$$

where

$$L_1 = L\cdot A, \qquad L_2 = L\cdot F. \tag{2'}$$

It is clear that $L_1\cdot D\neq 0$, and hence $L\cdot C\neq 0$, all points of A in D belonging to C. We now note that

$$\phi(C) \subset H, \tag{3}$$

since $\phi(C)\cdot\phi(B)=0$ and $\phi(C)\supset\phi(a)$. Also, that

$$\phi(L_1)\cdot H \subset \phi(C), \tag{4}$$

since

$$\phi(L_1)\cdot H \subset \phi(A)\cdot H \subset [\phi(B)+\phi(C)]\cdot H = \phi(C)\cdot H = \phi(C).$$

* That is, there exist, in $D-C$, two points P and Q which are not joined by any subcontinuum of $D-C$.

† See P. Alexandroff, *Sur la décomposition de l'espace par les ensembles fermés*, Comptes Rendus, vol. 184 (1927), pp. 425–428.

‡ That $B\neq 0$ is an immediate consequence of the Alexander Addition Theorem. (See proof of Theorem 2 below.)

Suppose that H contains a point which is not in $\phi(L_1)$. Then we can express H as the sum of two mutually exclusive sets, H_1 and H_2, where

$$H_1 = H \cdot \phi(L_1), \quad H_2 = H - H_1. \tag{5}$$

Since $\phi(L_1)$ is closed, it follows from (5) that H_1 is closed in H. Then, since H is connected, H_1 contains a limit point, $\phi(P)$, of H_2.

Let ϵ be a positive number less than $\rho[\phi(P), \phi(B)]$, and such that $F_{n-1}[\phi(P), \epsilon]$ contains a point, Q, of H, that is not in $\phi(L_1)$. Such a point exists, of course, since P is a limit point of the set of such points. Let

$$\phi(L_1) \cdot \bar{S}_{n-1}[\phi(P), \epsilon] = \phi(L_3), \tag{6}$$

$$\phi(L_1) \cdot \{E_{n-1} - S_{n-1}[\phi(P), \epsilon]\} = \phi(L_4). \tag{7}$$

That $\phi(L_3) \neq 0$ is obvious, since $\phi(P) \subset \phi(L_3)$. That $\phi(L_4) \neq 0$ follows from the following considerations. If we denote the set of points of L that are not in L_2 by L', then L_2 contains a limit point, x, of L'. Since $L' \subset C$, the point x is contained in $\overline{C} - C = B$. By a theorem of Miss Mullikin,* the continuum L contains a connected set, L'', which contains P and has at least one limit point in B, but contains no point of B. It is easy to see that $L'' \subset L_1$, and consequently that $\phi(L'')$ must have points in $F_{n-1}[\phi(P), \epsilon]$ that are also points of $\phi(L_1)$. Thus $\phi(L_4) \neq 0$.

Since $\phi(L_3) \cdot \phi(L_4) \subset F_{n-1}[\phi(P), \epsilon]$, and since Q is a point of $F_{n-1}[\phi(P), \epsilon]$ that is not in $\phi(L_1)$, it is clear that $\phi(L_3) \cdot \phi(L_4)$ does not fill up the surface $F_{n-1}[\phi(P), \epsilon]$, and consequently that the $(n-2)$th Betti number (mod 2)† of $\phi(L_3) \cdot \phi(L_4)$ is zero. In symbols,

$$p^{n-2}[\phi(L_3) \cdot \phi(L_4)] = 0. \tag{8}$$

By (2), (6) and (7),

$$L = L_1 + L_2 = L_3 + (L_2 + L_4). \tag{9}$$

---

* Anna Mullikin, *Certain theorems relating to plane connected point sets*, Transactions of this Society, vol. 24 (1922), pp. 144–162.

† See P. Alexandroff, *Une définition des nombres de Betti pour un ensemble fermé quelconque*, Comptes Rendus, vol. 184 (1927), pp. 317–319.

Neither of the sets L_3, L_2+L_4, is identical with L; for L_2+L_4 does not contain P, and L_3 contains no point of L_2. Consequently, since L is an *irreducible* cut between c and c_1,

$$(10) \qquad \begin{aligned} c + c_1 &\sim 0 \quad (\text{mod } 2, E_n - L_3), \\ c + c_1 &\sim 0 \quad [\text{mod } 2, E_n - (L_2 + L_4)]. \end{aligned}$$

Now

$$(11) \qquad L_3 \cdot (L_2 + L_4) = L_2 \cdot L_3 + L_3 \cdot L_4 = L_3 \cdot L_4.$$

Since $L_3 \cdot L_4$ is homeomorphic with $\phi(L_3) \cdot \phi(L_4)$, and the Betti number of a closed set is an analysis situs invariant, it follows from (8) and (11) that

$$(12) \qquad p^{n-2}[L_3 \cdot (L_2 + L_4)] = p^{n-2}(L_3 \cdot L_4) = 0.$$

Consequently, by virtue of Alexandroff's generalization of the Phragmén-Brouwer theorem,* and relations (9), (10), and (12),

$$(13) \qquad c + c_1 \sim 0 \quad (\text{mod } 2, E_n - L).$$

But this is a contradiction of the fact that L is a cut of E_n between c and c_1. Thus the supposition that H contains a point not in $\phi(L_1)$ leads to a contradiction, and

$$(14) \qquad H \subset \phi(L_1).$$

From relations (3) and (14) we have that

$$(15) \qquad \phi(C) \subset \phi(L_1),$$

and hence, from (3) and (15),

$$(16) \qquad \phi(C) \subset H \cdot \phi(L_1).$$

From relations (4), (14), and (16) it follows that

$$\phi(C) = H \cdot \phi(L_1) = H.$$

As H is an open connected subset of E_{n-1}, the theorem is proved, the relations $\overline{H} - H = \phi(\overline{C} - C) = \phi(B)$ being an immediate consequence of the fact that $H = \phi(C)$.

THEOREM 2. *In E_n, let D be a bounded domain such that all*

* P. Alexandroff, *Une généralization nouvelle du théorème de Phragmén-Brouwer*, Comptes Rendus, vol. 184 (1927), pp. 575–577.

1-cycles of D are homologous to zero in D. Then if two points c_1 and c_2 of D are separated in D by $A \cdot D$, there is a component, C, of $A \cdot D$ which separates c_1 and c_2 in D.*

PROOF. Let that component of $D-A \cdot D$ which contains c_1 be denoted by G. Then $\overline{G}-G$ is a cut of E_n between c_1 and c_2, and contains an irreducible cut, L, between c_1 and c_2. Let L_1 and L_2 be defined as in (2') above. Let P be a point of $L_1 \cdot D$, and let C be the component of $A \cdot D$ determined by P. Let $B = A \cdot F$, where F is the boundary of D, and let H be the component of $E_{n-1} - \phi(B)$ determined by $\phi(P)$. As before, if we suppose H contains a point which is not in $\phi(L_1)$, we can separate H according to relations (5), and proceed to a contradiction; consequently $H \subset \phi(L_1)$. Since $\phi(L_1) \subset \phi(A)$, we have $H \subset \phi(A)$ and accordingly $\phi^{-1}(H) \subset A$. Since we know that H contains no point of $\phi(B)$, $\phi^{-1}(H) \subset A-B$. Therefore $\phi^{-1}(H) \subset C$. It is clear that $\phi(C) \subset H$, and therefore $C \subset \phi^{-1}(H)$. Consequently $\phi^{-1}(H) = C$ and $H = \phi(C)$. Since, as noted above, $H \subset \phi(L_1)$, it follows that C is the component of $L_1 \cdot D$ determined by P. Thus every component of $L_1 \cdot D$ is homeomorphic with an $(n-1)$-dimensional domain of E_{n-1} whose boundary is in $\phi(B)$.

Let t be any 1-chain in D bounded by c_1+c_2. Not more than a finite number of the components of $L_1 \cdot D$ contain points of t. For suppose infinitely many contain points of t, and let C_1, C_2, C_3, $\cdots$ denote these components; they form a denumerable collection, since, as just shown, every component of $L_1 \cdot D$ is homeomorphic with a domain of E_{n-1}, and no two components have points in common. Let x_i be a point of $C_i \cdot t$, $(i = 1, 2, 3, \cdots)$. Then the set $\sum_{i=1}^{\infty} x_i$ has at least one limit point, y, on t. As L_1 is closed, $y \subset L_1$, and there is a component, U, of $L_1 \cdot D$ that contains y. But $\phi(y)$ is an interior point of the domain $\phi(U)$ and cannot be a limit point of the set $\sum_{i=1}^{\infty} \phi(x_i)$. The contradiction

* We refer here to modulo 2 homologies. See J. W. Alexander, *Combinatorial analysis situs*, Transactions of this Society, vol. 28 (1926), pp. 301–329. The necessity for this condition on the 1-cycles is made evident by the case where D is the interior of the anchor-ring in E_3; for a plane may be passed through D, in this case, in such a way that two points of D are separated by two components of the plane section, but not separated by either one of the components. An important case where the condition is satisfied is of course that in which D is bounded by the topological $(n-1)$-sphere.

is obvious. Let those components of $L_1 \cdot D$ that have points in common with t be denoted by $K_1, K_2, \cdots, K_m$.

One of the components K_i separates c_1 from c_2 in D. To show this, we note first that their sum $\sum_{i=1}^{m} K_i$ separates c_1 from c_2 in D. For suppose this is not the case. Then, denoting the set $L_1 \cdot D - \sum_{i=1}^{m} K_i$ by L_1', there exist 1-chains T_1^1 $(=t$, say) and T_2^1, such that*

$$(17)\quad \begin{cases} T_1^1 \to c_1 + c_2, [D - L_1']; [E_n - (L_2 + L_1')], \\ T_2^1 \to c_1 + c_2, \left[D - \sum_{i=1}^{m} K_i\right]; \left[E_n - \left(L_2 + \sum_{i=1}^{m} K_i\right)\right]. \end{cases}$$

The sets $L_2 + L_1'$ and $L_2 + \sum_{i=1}^{m} K_i$ are closed, and have in common only L_2. The latter set, however, lies entirely in F. Accordingly, by the hypothesis, there exists a 2-chain, T^2, bounded by $T_1^1 + T_2^1$ in D, and we have

$$(18)\quad T_1^1 + T_2^1 \sim 0, \left[E_n - (L_2 + L_1') \cdot \left(L_2 + \sum_{i=1}^{m} K_i\right)\right].$$

By Alexander's Addition Theorem,† and by (17) and (18),

$$(19)\qquad c_1 + c_2 \sim 0, \quad (E_n - L).$$

But (19) contradicts the fact that L is a cut of E_n between c_1 and c_2. Consequently one of the sets $L_1', \sum_{i=1}^{m} K_i$, separates c_1 and c_2 in D, and as the former set has no points in t, it is obvious that c_1 and c_2 are separated by $\sum_{i=1}^{m} K_i$ in D.

The proof can now be completed by a finite number of steps. If K_1 does not separate c_1 and c_2 in D, we can show by use of the Alexander Addition Theorem that $\sum_{i=1}^{m} K_i$ separates c_1 and c_2 in D. By process of elimination we must finally arrive at a set K_j, $(1 \leqq j \leqq m)$, which separates c_1 and c_2 in D. As $K_j \subset L_1$ and $L_1 \subset A$, the component $K_j \subset A$ and the theorem is proved.

THEOREM 3. *Let P be a point of A. Then for any positive number ρ there exists a positive number ϵ such that if Q is any point of*

* If C^i denote an i-cycle, then the relation $M^{i+1} \to C^i$ is to be interpreted "M^{i+1} is an $(i+1)$-chain bounded by C^i." See J. W. Alexander, loc. cit. All congruences and homologies used in the present instance are to be understood as modulo 2, without explicit statement of that fact in the relations given.

† J. W. Alexander, *A proof and extension of the Jordan-Brouwer separation theorem*, Transactions of this Society, vol. 23 (1922), pp. 333–349, Corollary W^5.

$(E_n-A)\cdot S_n(P, \epsilon)$, then that component of $(E_n-A)\cdot S_n(P, \rho)$ which contains Q has P as a boundary point.

PROOF. Suppose there exists a positive number ρ for which the theorem is not true. Let ϵ_1 be a positive number less than 1 as well as less than ρ. Then the neighborhood $S_n(P, \epsilon_1)$ contains a point, x_1, of E_n-A, such that the component, G_1, of $(E_n-A)\cdot S_n(P, \rho)$ which contains x_1 does not have P as a boundary point. Then $\overline{G}_1-G_1$ is a cut of E_n between x_1 and P. Accordingly, $(\overline{G}_1-G_1)\cdot A\cdot S_n(P, \rho)$ is a cut of $S_n(P, \rho)$ between x_1 and P, and by virtue of Theorem 2 (as applied to the closed set $(\overline{G}_1-G_1)\cdot A$) there is a component, C_1, of this set, which separates x_1 and P in $S_n(P, \rho)$. By Theorem 1 the set $\phi(C_1)$ is a domain of E_{n-1} whose boundary is $\phi(\overline{C}_1-C_1) \subset \phi[A\cdot F_n(P, \rho)]$. It is clear, then, that C_1 is a component of $A\cdot S_n(P, \rho)$.

Let ϵ_2 be a positive number less than $\frac{1}{2}$ as well as less than ϵ_1 and $\rho(P, C_1)$. Then $S_n(P, \epsilon_2)$ contains a point x_2 of E_n-A such that the component, G_2, of $(E_n-A)\cdot S_n(P, \rho)$ which contains x_2 does not have P as a boundary point. As before, there is a component, C_2, of $A\cdot S_n(P, \rho)$ which separates x_2 from P in $S_n(P, \rho)$.

Continuing in this way, we obtain a sequence of distinct points $x_1, x_2, x_3, \cdots$, having P as a sequential limit point, and a sequence $C_1, C_2, C_3, \cdots$ of distinct components of $A\cdot S_n(P, \rho)$ such that for every i, x_i and P are separated in $S_n(P, \rho)$ by C_i.

From the fact that $x_i \subset S_n(P, \epsilon_i)$, and C_i separates x_i and P in $S_n(P, \rho)$, it follows that there is a point y_i of C_i in $S_n(P, \epsilon_i)$. That the sequence $y_1, y_2, y_3, \cdots$ has P as a sequential limit point is obvious. Let

$$\overline{C}_i\cdot F_n(P, \rho) = B_i, \qquad (i = 1, 2, 3, \cdots). \tag{20}$$

Then $B_i \subset A$, and by Theorem 1, $\phi(C_i)$ is a domain of E_{n-1} whose boundary is $\phi(B_i)$.

For every i, $\phi(B_i)$ separates $\phi(y_i)$ from $\phi(P)$ in E_{n-1}, and hence $\phi(P)$ is a limit point in E_{n-1} of the set $\sum_{i=1}^{\infty}\phi(B_i)$. But then in E_n, P must be a limit point of the set $\sum_{i=1}^{\infty}B_i$, which is absurd since by (20) the sets B_i are all in $F_n(P, \rho)$.

The following theorem now follows simply from Theorem 3.

THEOREM 4. *Every point of A is regularly accessible from E_n-A.*

THE UNIVERSITY OF MICHIGAN

A PROPERTY OF CONTINUA SIMILAR TO LOCAL CONNECTIVITY*

BY W. A. WILSON

1. *Introduction.* An important property of locally connected compact spaces is that the complement of every continuum is the sum of a finite or enumerably infinite set of connected regions. If, however, a compact space Z is not locally connected and X is a sub-continuum, neither of the properties of Z being locally connected about X and $Z-X$ being the sum of an at most enumerably infinite set of connected regions involves the other. A property that is stronger than that of Z being locally connected about X is the following, introduced in a recent paper† by G. T. Whyburn: A sub-continuum X is ϵ-separated by a finite set of sub-continua of Z if, for each $\epsilon>0$, there is a finite set of sub-continua $\{F_i\}$ of Z such that $Z-\sum_1^m F_i$ is the sum of two separated sets Z_1 and Z_2, where Z_1 contains X and all points of $Z_1+\sum_1^m F_i$ have a distance less than ϵ from X. It is the purpose of this paper to describe another quasi-local property and its relation to those just mentioned, which in strength lies between local connectivity and ϵ-separability. The work will be carried out for separable metric continuous spaces in which the Bolzano-Weierstrass property is valid; such spaces will be called W-spaces for the sake of brevity.

DEFINITIONS. If a and b are disjoint bounded continua (or points) in a W-space Z, and Z can be expressed as the sum of two continua H and K such that $H\cdot b=K\cdot a=0$, we say that Z is *divisible* between a and b.

If Z is divisible between the bounded sub-continuum (or point) x and every sub-continuum of $Z-x$, then x is called *biregular*.

For example, let Z consist of a segment ab of unit length and an enumerable set of segments ab_n, each of unit length and inclined to ab at the angle π/n. Then Z is locally connected at a,

* Presented to the Society, September 9, 1930.

† *A generalized notion of accessibility*, Fundamenta Mathematicae, vol. 14, pp. 311–326.

or, as it is sometimes stated, a is a regular point, but a is not biregular, nor is it ϵ-separable, nor is its complement the sum of connected regions. However, ab has all these properties.

A condition weaker than biregularity is for Z to be divisible between x and every *point* of $Z-x$. Thus, if Z consists of a unit segment ab and an enumerable set of arcs $\{C_n\}$, each terminated by a and b and of radius n, and x is a point of $ab-(a+b)$, Z is divisible between x and every point of $Z-x$, but x is neither regular nor biregular, although its complement is a semi-continuum. Since in this example the point a is a biregular continuum, we see that biregularity is not, strictly speaking, a local property.

2. LEMMA. *Let Z be a connected W-space, x be a bounded sub-continuum, and Z be divisible between x and every point of $Z-x$. Then for each $\epsilon>0$, there is a least integer m such that the set of points of Z whose distances from x are not less than ϵ can be covered by m disjoint sub-continua of $Z-x$.*

PROOF. Here and elsewhere we denote the set of points of Z whose distances from x are less than ϵ by $V_\epsilon(x)$. If y is any point of $Z-V_\epsilon(x)$, we have $Z=H_y+K_y$, where H_y and K_y are continua and $y\cdot H_y=x\cdot K_y=0$. Let F be the frontier of $V_\epsilon(x)$. Then F is a bounded closed set and each of its points is an inner point of some K_y. By the Borel theorem, F is covered by a finite set of the continua $\{K_y\}$, say $K_1, K_2, \cdots, K_n$.

The set $T=Z-V_\epsilon(x)+\sum_1^n K_i$ is a closed set and each point of F is an inner point of some K_i. Hence F is covered by some finite set of components of T, say $C_1, C_2, \cdots, C_r$. Unless $r=1$, T is not connected; suppose that it can be decomposed into s mutually disjoint closed sets $\{T_i\}$. If $s>r$, some T_i, say T_s, would contain no C_i and Z would be the sum of two disjoint closed sets, T_s and $V_\epsilon(x)+\sum_1^{s-1}T_i$, a contradiction. Thus $s\leqq r$ and we can take $s=r$, since there are r components $\{C_i\}$. It follows, then, that each $T_i=C_i$ and so $Z-V_\epsilon(x)$ can be covered by r disjoint continua.

As r is finite, there is some least value of r which we take as m.

3. THEOREM 1. *Let Z be a connected W-space and x be a biregular continuum. Then Z is locally connected about x.*

In a previous note* it was shown that, if Z is divisible between every pair of disjoint sub-continua, then Z is everywhere locally connected. The demonstration of the above theorem is nearly the same.

By the lemma of the previous section we know that for each $\epsilon>0$ there is a least integer m such that $Z-V_\epsilon(x)$ is covered by m disjoint continua $\{S_i\}$, none of which contains any point of x. From this point on the proof in the reference given, beginning at the top of page 87, is applicable almost verbatim, since Z is divisible between x and every sub-continuum of $Z-x$.

This theorem together with the first example in §1 shows that the condition of biregularity is stronger than that of local connectivity. However, if Z is locally connected about the bounded continuum x and also Z is divisible between x and every *point* of $Z-x$, then x is biregular. For, if y is a sub-continuum of $Z-x$ and ϵ is less than the distance between x and y, there is a continuum X such that some $V_\delta(x)\subset X$ and $X\subset V_\epsilon(x)$. By §2, $Z-V_\delta(x)$ is covered by m disjoint continua $\{S_i\}$, none of which meets x. If $y\subset S_1$, then $X'=X+\sum_2^m S_i$ is a continuum, $Z=X'+S_1$, $y\cdot X'=0$, and $x\cdot S_1=0$. Hence x is biregular.

4. THEOREM 2. *Let Z be a connected W-space, x be a bounded sub-continuum, and Z be divisible between x and every point of $Z-x$. Then $Z-x$ has a finite or enumerable set of components and each component is both a region and a semi-continuum.*

PROOF. Take $\epsilon_1>0$. By §2 we know that $Z-V_{\epsilon_1}(x)$ is covered by m_1, but not less than m_1, mutually disjoint continua $\{K_r^1\}$, $r=1, 2, \cdots, m_1$, no one of which contains a point of x. Take ϵ_2 less than half ϵ_1 and less than the least distance between x and any K_r^1. Then $Z-V_{\epsilon_2}(x)$ is covered by m_2, but not less than m_2, mutually disjoint continua $\{K_r^2\}$, $r=1, 2, \cdots, m_2$, no one of which contains a point of x. As $K_1^1\subset Z-V_{\epsilon_2}(x)$, K_1^1 is an inner set of some K_r^2, say K_1^2. No $K_r^1\subset K_1^2$ if $r\neq 1$, for then $Z-V_{\epsilon_1}(x)$ would be covered by m_1-1 continua. Let, then, $K_1^1\subset K_1^2$, $K_2^1\subset K_2^2$, $K_3^1\subset K_3^2$, etc. We note that $m_1\leqq m_2$.

Now take ϵ_3 less than half ϵ_2 and less than the least distance between x and any K_r^2 and define the continua $\{K_r^3\}$, $r=1$,

* *A property of continua equivalent to local connectivity*, this Bulletin, vol. 36, pp. 85–88.

$2, \cdots, m_3$, as before. Continue this process indefinitely. We then have

$$K_1^1 \subset K_1^2 \subset K_1^3 \subset \cdots,$$

$$K_2^1 \subset K_2^2 \subset K_2^3 \subset \cdots,$$

$$K_3^1 \subset K_3^2 \subset K_3^3 \subset \cdots.$$

For each r let L_r be the union of the sequence of continua $\{K_r^i\}$. Then each L_r is a semi-continuum and, since $\epsilon_i \to 0$, $Z-x$ is the union of the semi-continua $\{L_r\}$. If, from some i on, we have $m_i = n$, there are n semi-continua $\{L_r\}$; if $m_i \to \infty$ as $\epsilon_i \to 0$, there is an enumerable set of the semi-continua $\{L_r\}$.

We first note that $L_r \cdot L_s = 0$, if $r \neq s$. For in the case of every i, $K_r^i \cdot K_s^i = 0$ if $r \neq s$. We now show that each L_r is a region. If y is any point of L_r, $Z = H + K$, where H and K are continua and $x \cdot K = y \cdot H = 0$. Then for some i, $K \cdot V_{\epsilon_i}(x) = 0$, whence K lies in some K_r^i; that is, $K \subset L_r$. But y is an inner point of K, and consequently of L_r. Finally, since the sets $\{L_r\}$ are disjoint connected regions, they are components of $Z-x$. Thus the theorem is proved.

NOTE. The theorem is valid a fortiori if x is biregular. The proof also shows that each component of $Z-x$ is the union of an ascending sequence of continua, each of which is an inner set of those following. The converse is not true, as is seen from the first example in §1, but we do have the following as a sort of converse.

THEOREM 3. *Let Z be a connected compact space, x be a sub-continuum, Z be locally connected about x, and every component of $Z-x$ be a region and a semi-continuum which is the union of an ascending sequence of continua such that each one is an inner set of the following. Then x is biregular.*

PROOF. Let R be any component of $Z-x$, S be the sum of the others, and y be a point of R. Then y is an inner point of some sub-continuum K of R. Let ϵ be half the distance between x and K. Since Z is locally connected about x, there is a $\delta > 0$ and a continuum X such that $V_\delta(x) \subset X \subset V_\epsilon(x)$ and $K \cdot X = 0$.

Every component of $\overline{R} - (X+K)$ has limiting points on X or K, but none on x or y. Let L and M be the sum of X and K, respectively, and the components of $\overline{R} - (X+K)$ having limiting points on X and K, respectively. Then $\overline{L}+S$ and $\overline{M}$ are

continua and $Z=(\overline{L}+S)+\overline{M}$ is a division of Z between x and y. Since Z is locally connected about x and is divisible between x and every point of $Z-x$, x is biregular.

Simple examples show that the theorem may fail if any of the hypotheses are omitted.

5. THEOREM 4. *Let Z be a connected compact space, x be a biregular sub-continuum, and R_i be any component of $Z-x$. If $M_i=x+R_i$, each M_i and the union of any number of the M_i are biregular continua. Also x is biregular with respect to each M_i.*

PROOF. Let A be the union of any number of the continua $\{M_i\}$ and let y be any sub-continuum of $Z-A$. Now A is a continuum, because it is connected and $Z-A$ is the sum of a finite or enumerable set of disjoint regions. The continuum y lies in some R_i, say in R_1, and $Z=H+K$, where H and K are continua and $x\cdot K=y\cdot H=0$. As $x\cdot K=0$, $K\subset R_1$ and $A\subset H$; hence $A\cdot K=y\cdot H=0$. Thus A is biregular.

Now consider any one of the continua $\{M_i\}$, say M_1. If y is a sub-continuum of M_1-x, it lies in R_1. Then $K\subset R_1$ and $x\subset H$. As x is bounded, each component of $H-x$ has limiting points on x; hence $H_1=H\cdot M_1=x+H\cdot R_1$ is a continuum. Thus M_1 is the union of two continua K and H_1, $y\cdot H_1=0$, and $x\cdot K=0$, which was to be proved.

6. THEOREM 5. *Let Z be a connected W-space and x a biregular sub-continuum. Then for every $\epsilon>0$ there is a least integer m such that Z is the union of a continuum X containing some $V_\delta(x)$ and contained in $V_\epsilon(x)$, and m mutually disjoint continua $\{K_i\}$, each of which contains no point of x.*

PROOF. By §3 there is a continuum X such that some $V_\delta(x)\subset X\subset V_\epsilon(x)$. By §2 there are m mutually disjoint continua $\{K_i\}$ such that $Z-V_\delta(x)\subset\sum_1^m K_i$ and $x\cdot K_i=0$ for each value of i. As m is finite, there is clearly some choice of X and the K_i so that m has a least value.

COROLLARY. *If in addition to the above hypotheses Z is at most n-coherent,* then x is ϵ-separable by a finite set of continua.*

* This term signifies that for every partition of Z into two continua H and K, $H\cdot K$ has not more than n components. This mode of classifying continua is a modification of a definition by C. Kuratowski.

For $X \cdot K_i$ has at most n components for each value of i and so $X \cdot \sum_1^m K_i$ has at most mn components.

As remarked in §4, the integer m may increase indefinitely as $\epsilon \to 0$, or it may remain bounded. Generalizing Menger's notion of order of a point, we may say in the former case that x is of order w; in the latter, of finite order. If Z is unicoherent, the order of x is w if $Z-x$ has an infinite set of components, and is the same as the number of components of $Z-x$ if that number is finite. This corollary shows that for at most n-coherent continua biregularity is a local property.

7. THEOREM 6. *Let Z be a connected W-space, x be a bounded sub-continuum, and x be ϵ-separated by a finite set of continua. Then x is biregular.*

PROOF. Let y be any sub-continuum of $Z-x$ and ϵ be half the distance between x and y. By hypothesis there is a region R such that $x \subset R \subset V_\epsilon(x)$ and F, the frontier of R, is contained in the sum of m bounded sub-continua $\{F_i\}$ of $V_\epsilon(x)$.

Let $G=Z-R$. The closed set G cannot be decomposed into more than m mutually disjoint closed sets $\{G_i\}$, for, if this were possible, one of them, say G_1, would contain no point of F and we would have Z decomposed into two disjoint closed sets, G_1 and $Z-G_1$. Hence G is the sum of r mutually disjoint continua $\{G_i\}$, and $r \leqq m$.

Similarly, since every component of R has limiting points on at least one F_i, $\overline{R}$ can be decomposed into k mutually disjoint continua $\{C_i\}$, where $k \leqq m$, and one of them, say C_1, contains some $V_\delta(x)$. Clearly $G+\sum_2^k C_i$ is the sum of a finite number of mutually disjoint continua $\{K_j\}$, none of which contains a point of x, all of which meet C_1, and one of which, say K_1, contains y. Let H_1 be the union of C_1 and the other continua $\{K_j\}$. Then $Z=H_1+K_1$, $y \cdot H_1=0$, and $x \cdot K_1=0$. Hence the theorem is proved.

8. *Compact Space.* If Z is a bounded W-space, that is, a compact space, Z itself formally satisfies the definition of biregularity. It is easily seen from the definition that the divisor of a descending sequence of biregular continua is biregular; hence each point x lies on a continuum X irreducible with respect to the properties of being biregular and containing x.

If there is an upper semi-continuous decomposition of Z into sub-continua about each of which Z is locally connected, it is clear that each element is biregular. It would then appear that at least in the case of at most n-coherent continua the irreducibly biregular sub-continua would serve as elements of such a decomposition. Unfortunately they need not be unique. However, one special case is not without interest.

THEOREM 7. *Let Z be a connected compact space and let Z be divisible between no two points of any irreducibly biregular sub-continuum. Then there is an upper semi-continuous decomposition of Z into elements $\{X\}$, which are the irreducibly biregular sub-continua of Z.*

PROOF. We first note that each point of Z lies on only one irreducibly biregular sub-continuum. For, if a point x were on two such, say X and X', and y were a point of X' not on X, a division of Z between X and y would be a division between two points of X'.

Now let $\{X_i\}$ be a sequence of elements and X be an element such that, for $\epsilon>0$ and every i larger than some i', $X_i \cdot V_\epsilon(X) \neq 0$. If $Z-X$ contains a point of the upper closed limiting set X' of the sequence $\{X_i\}$, we can assume that for a partial sequence each X_i contains two points x_i and y_i such that $x_i \to a$ and $y_i \to b$, where a lies in X and b in $Z-X$. As Z is divisible between X and b, and so between a and b, there is an i so large that Z is divisible between x_i and y_i, a contradiction. Since Z is locally connected about each X, the theorem is proved.

In this special case the decomposition is the finest possible, for some elements of a finer decomposition would be biregular sub-continua and proper sub-continua of some X, an impossibility. If there is a region composed of biregular sub-continua $\{X\}$ such that Z is divisible between no two points of any X, we can obtain an upper semi-continuous decomposition analogous to that into prime parts by taking the components of the remainder of Z as elements.

YALE UNIVERSITY

ON THE WEDDERBURN NORM CONDITION FOR CYCLIC ALGEBRAS*

BY A. A. ALBERT

1. *Introduction.* Let F be any non-modular field, i a root of a cyclic equation in F of degree n and with roots $\theta^r(i)$. Suppose that A is a cyclic algebra with basis

$$i^r y^s, \qquad (r, s = 0, 1, \cdots, n-1),$$

where

$$y^r i = \theta^r(i) y^r, \; y^n = \gamma \text{ in } F.$$

J. H. M. Wedderburn has proved† that A is a division algebra if γ^r is not the norm, $N(a)$, of any a in $F(i)$ for every positive integer r less than n. It has never been shown, however, that this condition is a necessary one; but the problem of finding complete necessary and sufficient conditions has been reduced to the case n a power of a single prime. ‡

In the present paper cyclic algebras of order sixteen with the corresponding cyclic quartic in its canonical form§

$$\phi(\omega) \equiv \omega^4 + 2\nu(1+\Delta^2)\omega^2 + \nu^2\Delta^2(1+\Delta^2) = 0$$

such that ν and Δ are in F, and $\tau = 1+\Delta^2$ is not the square of any quantity of F, are considered. The norm $N(a)$ of a polynomial in i is a rather complicated quartic form in four variables, yet we can secure the result that $\gamma^2 = N(a)$ if and only if $\gamma = \alpha^2 - \beta^2\tau$ for α and β in F, a curious property of cyclic quartic fields. When the above equation is satisfied the algebra A is expressible as a direct product of two generalized quaternion algebras. Necessary and sufficient conditions are secured that our algebras A of order sixteen be division algebras, and it is shown that for the particularly interesting case where F is *the field of all rational numbers* the Wedderburn condition is *necessary as well as sufficient.*

---

* Presented to the Society, December 30, 1930.

† Transactions of this Society, vol. 15 (1914), pp. 162–166.

‡ See a paper by the author, *On direct products, cyclic algebras, and pure Riemann matrices*, to appear in the Transactions of this Society, January, 1931.

§ See R. Garver, *Quartic equations with certain groups*, Annals of Mathematics, vol. 29 (1928), pp. 47–51.

2. *The Basic Theorem.* Let $F(x)$ be a cyclic quartic field. Then it is known (loc. cit.) that $F(x)=F(i)$, where i satisfies the equation

$$\phi(\omega) \equiv \omega^4 + 2\nu\tau\omega^2 + \nu^2\Delta^2\tau = 0, \tag{1}$$

with $\tau=1+\Delta^2$ not the square of any quantity of F and

$$\nu \neq 0,\ \tau,\ \Delta \neq 0 \tag{2}$$

all in F. Moreover if we define u by the equation

$$i^2 = \nu(u-\tau), \tag{3}$$

then

$$u^2 = \tau, \qquad \theta(i) = \frac{i}{\Delta}(u+1), \tag{4}$$

is the polynomial whose iteratives $i=\theta^0(i)=\theta^4(i), \theta(i), \theta^2(i)=-i$, $\theta^3(i)=\theta(-i)=-\theta(i)$ give the four roots in $F(i)$ of $\phi(\omega)=0$. Every quantity of $F(i)$ is expressible in the form

$$a = a_1 + a_2 i, \qquad (a_1 \text{ and } a_2 \text{ in } F(u)), \tag{5}$$

and $a=0$ if and only if $a_1=a_2=0$. A quantity

$$a_1 = \alpha_1 + \alpha_2 u, \qquad (\alpha_1 \text{ and } \alpha_2 \text{ in } F), \tag{6}$$

is zero if and only if $\alpha_1=\alpha_2=0$; and similarly

$$\alpha_1^2 - \alpha_2^2\tau \tag{7}$$

vanishes if and only if $\alpha_1=\alpha_2=0$ by our restriction on τ.

We shall use repeatedly the following simple lemma.

LEMMA 1. *Every product of a finite number of scalars of the forms*

$$\lambda^2 - \mu^2\tau, \tag{8}$$

$$(\lambda^2 - \mu^2\tau)^{-1}, \qquad \lambda^2 - \mu^2\tau \neq 0, \tag{8'}$$

with λ and μ in F, is expressible in the form (8) *for λ and μ in F.*

The truth of this is evident since

$$(\lambda_1 + \mu_1 u)(\lambda_2 + \mu_2 u) = (\lambda_1\mu_1 + \lambda_2\mu_2\tau) + (\lambda_1\mu_2 + \lambda_2\mu_1)u,$$

and hence

(9) $(\lambda_1^2 - \mu_1^2\tau)(\lambda_2^2 - \mu_2^2\tau) = (\lambda_1\mu_1 + \lambda_2\mu_2\tau)^2 - (\lambda_1\mu_2 + \lambda_2\mu_1)^2\tau;$

while if $\epsilon = \lambda^2 - \mu^2\tau \neq 0$, then

(10) $\epsilon^{-1} = \epsilon^{-2}\epsilon = \epsilon^{-2}(\lambda^2 - \mu^2\tau) = (\lambda\epsilon^{-1})^2 - (\mu\epsilon^{-1})^2\tau.$

Let us now assume that $\gamma \neq 0$ is a scalar in F, such that $\gamma^2 = N(a)$, where a is in the cyclic field $F(i)$. We may write $a^{(r)} = a[\theta^r(i)]$, $(r = 0, 1, \cdots)$, whence $a'' = a(-i)$. Then $u' = -u$; and if a_1 is in $F(u)$ so that a_1 has the form $a_1 = \alpha_1 + \alpha_2 u$, we have $a_1 = a_1''$ and

(11) $N(a_1) = a_1 a_1' a_1'' a_1''' = (a_1 a_1')^2 = (\alpha_1^2 - \alpha_2^2\tau)^2.$

Let us write $\gamma^2 = N(a)$, where

(12) $a = a_2 + a_3 i, \qquad (a_2 \text{ and } a_3 \text{ in } F(u)).$

We shall first consider the case $a_3 = 0$. Then $a = c_2 = \alpha_3 + \alpha_4 u$, and

(13) $\gamma^2 = (\alpha_3^2 - \alpha_4^2\tau)^2.$

This equation in a field F implies that

(14) $\gamma = \pm (\alpha_3^2 - \alpha_4^2\tau).$

If $\gamma = \alpha_3^2 - \alpha_4^2\tau$, we have expressed γ in the form

(15) $\gamma = \alpha^2 - \beta^2\tau$

with α and β in F, the result desired. Since $\tau = 1 + \Delta^2$, we have

(16) $-1 = \Delta^2 - \tau.$

Hence if $\gamma = -(\alpha_3^2 - \alpha_4^2\tau)$, then $\gamma = (\Delta^2 - \tau)(\alpha_3^2 - \alpha_4^2\tau)$; and, by Lemma 1, γ has again the desired form (15).

Next let $a_3 \neq 0$. Then, if $a_3 = \lambda_3 + \lambda_4 u$, $a_1 = a_3^{-1}a_2$, we have

(17) $N(a) = N[a_3(a_1 + i)] = (\lambda_3^2 - \lambda_4^2\tau)^2 N(a_1 + i).$

Let $\delta = \gamma(\lambda_3^2 - \lambda_4^2\tau)^{-1}$. Then

(18) $\delta^2 = \gamma^2(\lambda_3^2 - \lambda_4^2\tau)^{-2} = N(a_1 + i).$

But if $b = a_1 + i$, then $\delta^2 = (bb'')(bb'')'$ so that if $w = \delta[(bb'')']^{-1}$, then $\delta = \delta' = w'bb''$. It follows that $\delta^2 = w\,w'N(b) = w\,w'\delta^2$. Hence

(19) $ww' = 1, \quad w = bb''\delta^{-1}, \quad bb'' = \delta w,$

where $w=bb''\delta^{-1}=\xi_1+\xi_2 u$ is in $F(u)$. If $a_1=\alpha_1+\alpha_2 u$, α_1 and α_2 in F, we have, by (3),

$$bb'' = a_1^2 - i^2 = \alpha_1^2 + \alpha_2^2\tau + 2\alpha_1\alpha_2 u - \nu(u-\tau)$$
$$= (\alpha_1^2 + \alpha_2^2\tau + \nu\tau) + (2\alpha_1\alpha_2 - \nu)u.$$

From the linear independence of 1 and u this implies

$$\alpha_1^2 + \alpha_2^2\tau + \nu\tau = \delta\xi_1, \quad 2\alpha_1\alpha_2 - \nu = \delta\xi_2. \tag{20}$$

We obtain $2\alpha_1\alpha_2\tau - \nu\tau = \delta\xi_2\tau$, and by addition

$$\alpha_1^2 + 2\alpha_1\alpha_2\tau + \alpha_2^2\tau = \delta(\xi_1 + \xi_2\tau). \tag{21}$$

Since $1-\tau=-\Delta^2$, if we complete the square in (21), it becomes

$$(\alpha_1 + \alpha_2\tau)^2 + \alpha_2^2(\tau - \tau^2) = (\alpha_1 + \alpha_2\tau)^2 - (\alpha_2\Delta)^2\tau \tag{22}$$
$$= \delta(\xi_1 + \xi_2\tau).$$

Consider now the equation $ww'=1$, or

$$\xi_1^2 - \xi_2^2\tau = 1, \quad \xi_2^2\tau = (\xi_1+1)(\xi_1-1). \tag{23}$$

Let $\xi_1-1=2\pi$, $\xi_1+1=2\sigma$. Then

$$4\sigma\pi = \xi_2^2\tau. \tag{24}$$

Suppose first that $\xi_1+1=0$ so that $\sigma=0$ and $\xi_2=0$. Then $\xi_1+\xi_2\tau=\xi_1=-1=\Delta_1^2-\tau$. Hence in this case we have

$$\xi_1 + \xi_2\tau = \lambda_5^2 - \lambda_6^2\tau, \qquad (\lambda_5 \text{ and } \lambda_6 \text{ in } F). \tag{25}$$

Next let $\xi_1+1\neq 0$, so that $\sigma\neq 0$; and let us define ϵ by the equation

$$2\sigma\epsilon = \xi_2. \tag{26}$$

Then (24) gives $4\sigma\pi=4\sigma^2\epsilon^2\tau$, whence

$$\pi = \epsilon^2\sigma\tau. \tag{27}$$

But $2(\sigma-\pi)=\xi_1+1-(\xi_1-1)=2$, whence

$$1 = \sigma - \pi = \sigma - \epsilon^2\sigma\tau = \sigma(1 - \epsilon^2\tau). \tag{28}$$

Since $1-\epsilon^2\tau\neq 0$, using Lemma 1, we have

$$\sigma = \beta_1^2 - \beta_2^2\tau, \qquad (\beta_1 \text{ and } \beta_2 \text{ in } F), \tag{29}$$

so that $\xi_1=\pi+\sigma=\sigma(1+\epsilon^2\tau)$, and

$$\begin{aligned}\xi_1 + \xi_2\tau &= \sigma[(1+\epsilon^2\tau) + 2\epsilon\tau] = \sigma[(1+\epsilon\tau)^2 - (\epsilon\Delta)^2\tau] \\ &= (\beta_1^2 - \beta_2^2\tau)[(1+\epsilon\tau)^2 - (\epsilon\Delta)^2\tau] = \lambda_5^2 - \lambda_6^2\tau,\end{aligned} \tag{30}$$

for λ_5 and λ_6 in F, by Lemma 1. Hence *in all cases* (25) *is satisfied.*

If we now put $\beta_3=\alpha_1+\tau\alpha_2$, $\beta_4=\Delta\alpha_2$, (22) becomes

$$\delta(\lambda_5^2 - \lambda_6^2\tau) = \beta_3^2 - \beta_4^2\tau. \tag{31}$$

Suppose first that $\beta_3^2-\beta_4^2\tau=0$, whence $\beta_3=\beta_4=0$. Then our definitions above of β_3 and β_4 evidently give $\alpha_1=\alpha_2=0$, and (20) take the form $\nu\tau=\delta\xi_1$, $-\nu=\delta\xi_2$. Squaring each side of both these, we may write $\nu^2\tau^2=\delta^2\xi_1^2$, $\nu^2\tau=\delta^2\xi_2^2\tau$, whence, by subtraction and the use of the relations $1=\xi_1^2-\xi_2^2\tau$, $\tau=1+\Delta^2$, we obtain

$$\nu^2\tau^2 - \nu^2\tau = \tau(\nu^2\Delta^2) = \delta^2(\xi_1^2 - \xi_2^2\tau) = \delta^2. \tag{32}$$

Then $\tau=(\delta\nu^{-1}\Delta^{-1})^2$, which is a contradiction since τ is not the square of any quantity of F. Hence $\beta_3^2-\beta_4^2\tau\neq 0$. Thus $\lambda_5^2-\lambda_6^2\tau\neq 0$ has an inverse in F which has the form $\lambda_7^2-\lambda_8^2\tau$ by Lemma 1, and we may write

$$\begin{aligned}\gamma &= \delta(\lambda_3^2 - \lambda_4^2\tau) = (\lambda_3^2 - \lambda_4^2\tau)(\lambda_7^2 - \lambda_8^2\tau)(\beta_3^2 - \beta_4^2\tau) \\ &= (\alpha^2 - \beta^2\tau),\end{aligned} \tag{33}$$

again using Lemma 1. We have proved in all cases the first part of the following statement.

THEOREM 1. *A scalar $\gamma\neq 0$ in F has the property*

$$\gamma^2 = N(a) \tag{34}$$

for a in $F(i)$, a cyclic quartic field, if and only if

$$\gamma = \alpha^2 - \beta^2\tau, \qquad (\alpha \text{ and } \beta \text{ in } F), \tag{35}$$

where $F(u)$ is the quadratic subfield of $F(i)$ defined by (1) *and* (3), *and $u^2=\tau$.*

Moreover, when $\gamma=\alpha^2-\beta^2\tau$, we have $N(\alpha+\beta\mu)=(\alpha^2-\beta^2\tau)^2=\gamma^2$, which is the converse in the preceding theorem.

Suppose now that $\gamma=\alpha^2-\beta^2\tau$ and $\gamma^2=N(b)$. If $a=\alpha+\beta u$, so that $\gamma=aa'$, we have $\gamma^2=N(a)=N(b)$. It follows that $b\neq 0$ and $N(ab^{-1})=1$, $a=wb$, where $N(w)=1$. Thus we have the following corollary.

COROLLARY 1. *Let $\gamma=aa'$, where a is in $F(u)$. Then $\gamma^2=N(b)$ for b in $F(i)$ if and only if b is the product of a by a unit of $F(i)$.*

Since $-1=dd'$, where d is given in (16) and is in $F(u)$, we have also the following result.

COROLLARY 2. *The scalar $\gamma^2=N(b)$ for b in $F(i)$ if and only if $-\gamma=ee'$ for e in $F(u)$.*

3. *The Wedderburn Norm Condition.* For a cyclic algebra of order sixteen Wedderburn's condition becomes

$$\gamma^r \neq N(a), \qquad (r = 1, 2, 3).$$

It is easily shown* that if γ or γ^3 were a norm then A would not be a division algebra. Hence the only possible case is $\gamma^2=N(a)$. By Theorem 1 this implies that $\gamma=\alpha^2-\beta^2\tau$. Consider the sub-algebra

$$\sum = (y^r, uy^r), \qquad (r = 0, 1, 2, 3),$$

an algebra of order eight with $yu=-uy$, $y^4=\gamma$, $u^2=\tau$ in F, $y^2u=uy^2$. We shall write

$$(36) \qquad s = (e + y^2)y, \quad t = i(a_1 + y^2), \quad a_1 = \beta_1 q - \beta_2 u,$$

where we have used Corollary 2 to write

$$(37) \qquad -\gamma = ee', \quad e = \beta_1 + \beta_2 u, \qquad (\beta_1 \text{ and } \beta_2 \text{ in } F),$$

and have

$$(38) \qquad yi = \theta(i)y, \quad \theta(i) = qi, \quad q = \Delta^{-1}(u + 1), \quad qq' = -1,$$

since $\Delta^2 qq'=(u+1)(-u+1)=-(1+\Delta^2)+1=-\Delta^2$. We shall compute

$$\begin{aligned} st &= [(e + y^2)y][i(a_1 + y^2)] = (e + y^2)qi(a_1' + y^2)y \\ &= iq(e - y^2)(a_1' + y^2)y = iq[(ea' - \gamma) + (e - a_1')qy^2]y, \end{aligned}$$

since $ya=a'y$ for every a of $F(i)$ and $y^2i=-iy^2$. Now

* See the author's paper *On direct products, etc.*, loc. cit.

$$q(ea' - \gamma) = q(ea' + ee') = eq(a_1' + e')$$
$$= eq[\beta_1 q' + \beta_2 u + \beta_1 - \beta_2 u] = \beta_1 e[qq' + q] = \beta_1 e(q - 1).$$

Also

$$q(e - a_1') = q[(\beta_1 + \beta_2 u) - (\beta_1 q' + \beta_2 u)] = \beta_1(q + 1).$$

It follows that

$$st = \beta_1 i[e(q - 1) + (q + 1)y^2]y. \tag{39}$$

We have similarly

$$ts = [i(a_1 + y^2)][(e + y^2)y] = i[(a_1 e + \gamma) + (a_1 + e)y^2]y,$$
$$a_1 e + \gamma = a_1 e - ee' = e[(\beta_1 q - \beta_2 u) - (\beta_1 - \beta_2 u)] = \beta_1 e(q - 1),$$

while $a_1 + e = \beta_1 q - \beta_2 u + \beta_1 + \beta_2 u = \beta_1(q+1)$. We then obtain immediately from (39)

$$st = ts. \tag{40}$$

Consider the linear sets

$$B = (1, u, s, us), \qquad C = (1, y^2, t, y^2 t), \tag{41}$$

over F. We have the relations

$$su = -us, ty^2 = -y^2 t; uy^2 = y^2 u, ut = tu, sy^2 = y^2 s, st = ts, \tag{42}$$

so that every quantity of B is commutative with every quantity of C. We now show that

$$s^2 = (e + y^2)(e' + y^2)y^2 = [(ee' + \gamma) + (e + e')y^2]y^2 = 2\beta_1\gamma, \tag{43}$$

since $e + e' = 2\beta_1$, $ee' = -\gamma$. Also

$$t^2 = i^2(a_1 - y^2)(a_1 + y^2) = i^2(a_1^2 - \gamma)$$
$$= i^2[\beta_1^2 q^2 - 2\beta_1\beta_2 qu + \beta_2^2\tau + \beta_1^2 - \beta_2^2\tau]$$
$$= i^2\beta_1^2(q^2 + 1) - 2\beta_1\beta_2 qui^2,$$

since $\gamma = -ee'$. We have also $i^2 = \nu(u - \tau)$, so that

$$i^2 q = \nu\Delta^{-1}(u - \tau)\ (u + 1) = \Delta^{-1}\nu(\tau - \tau + u - u\tau)$$
$$= \Delta^{-1}\nu u(1 - \tau) = \Delta^{-1}\nu u(-\Delta^2) = -\Delta\nu u.$$

Moreover, we know that

$$i^2(q^2 + 1) = i^2[\Delta^{-2}(\tau + 2u + 1) + 1] = \Delta^{-2}i^2[2u + \tau + (1 + \Delta^2)]$$
$$= 2\nu\Delta^{-2}(u - \tau)(u + \tau) = 2\nu\Delta^{-2}(\tau - \tau^2) = -2\nu\tau.$$

Hence

$$t^2 = 2\nu\tau\beta_1(\beta_2\Delta - \beta_1). \tag{44}$$

We shall assume at this point that

$$\beta_1 \neq 0, \ \beta_2\Delta - \beta_1 \neq 0, \tag{45}$$

for otherwise either $s^2 = 0$ or $t^2 = 0$, and A is evidently not a division algebra since from their form neither s nor t is zero. As a consequence, $a_1^2 - \gamma \neq 0$ has an inverse in $F(u)$ and $e^2 - \gamma \neq 0$ has an inverse in $F(u)$ since $t^2 = i^2(a_1^2 - \gamma) \neq 0$, while if

$$e^2 - \gamma = e^2 - ee' = 0,$$

then $e(e - e') = 2\beta_1 e = 0$, contrary to the hypothesis that $\beta_1 \neq 0$, so that $e \neq 0$ has an inverse in $F(u)$. The sets B and C are generalized quaternion algebras over F, since in B

$$u^2 = \tau, \ s^2 = 2\beta_1\gamma, \ su = -us,$$

while in C

$$(y^2)^2 = \gamma, \ t^2 = 2\nu\tau\beta_1(\beta_2\Delta - \beta_1), \ y^2t = -ty^2,$$

and evidently from the form of s and t the quantities $1, u, s, us$ are linearly independent in F, and the quantities $1, y^2, t, y^2t$ are linearly independent in F, when $i^\alpha y^\beta$ $(\alpha, \beta = 0, 1, 2, 3)$ form a basis of A. The linear set $BC = CB$ of all sums of all products of quantities of B and quantities of C is an algebra, since a product

$$\left(\sum_\lambda b_{1\lambda}c_{1\lambda}\right)\left(\sum_\mu b_{2\mu}c_{2\mu}\right) = \sum_{\lambda,\mu}(b_{1\lambda}b_{2\mu})(c_{1\lambda}c_{2\mu})$$

is in BC because for every λ and μ the quantities $b_{1\lambda}b_{2\mu}$ are in B and $C_{1\lambda}C_{2\mu}$ are in C. Now BC contains $F(u)$ and hence $(\gamma - a_1^2)^{-1}$, $(\gamma - e^2)^{-1}$. Since BC contains s, t, e, a_1, y^2, and is an algebra, it contains

$$\begin{aligned}(\gamma - a_1^2)^{-1}(ty^2 - a_1t) &= (\gamma - a_1^2)^{-1}i(a_1y^2 + \gamma - a_1^2 - a_1y^2) \\ &= (\gamma - a_1^2)^{-1}(\gamma - a_1^2)i = i,\end{aligned}$$

and

$$\begin{aligned}(\gamma - e^2)^{-1}(y^2s - es) &= (\gamma - e^2)^{-1}[(y^2e + \gamma) - ey^2 - e^2]y \\ &= (\gamma - e^2)^{-1}(\gamma - e^2)y = y.\end{aligned}$$

But then BC contains the basis of A and has order sixteen. It follows that A is the *direct product* of B and C.

THEOREM 2. *Let $\gamma^2 = N(a)$ for some a of $F(i)$ so that $-\gamma = ee'$, where $e = \beta_1 + \beta_2 u$ and β_1 and β_2 are in F. Let $\beta_1 \neq 0$, $\beta_2\Delta \neq \beta_1$, a set of necessary conditions that A be a division algebra. Then the cyclic algebra A is the direct product of two generalized quaternion algebras $B = (1, u, s, us)$, $C = (1, j, t, jt)$, with $y^2 = j$, $su = -us$, $tj = -jt$, and*

$$(46)\quad u^2 = \tau,\ s^2 = 2\beta_1\gamma = \sigma,\ j^2 = \gamma,\ t^2 = \rho = 2\nu\beta_1\tau(\beta_2\Delta - \beta_1).$$

Consider now the direct product of *any two generalized quaternion algebras B and C.* It is known that d in B has the property that d^2 is in F if and only if

$$(47)\quad d = \lambda_1 u + \lambda_2 s + \lambda_3 us,\ d^2 = Q_1 = \lambda_1^2\tau + \lambda_2^2\sigma - \lambda_3^2\sigma\tau,$$

with λ_1, λ_2 and λ_3 in F. Similarly if f is in C then f^2 is in F if and only if

$$(48)\quad f = \lambda_4 j + \lambda_5 t + \lambda_6 jt,\ f^2 = Q_2 = \lambda_4^2\gamma + \lambda_5^2\rho - \lambda_6^2\gamma\rho$$

for λ_4, λ_5, and λ_6 in F. *Suppose first that $Q \equiv Q_1 - Q_2$ is a null form,* that is, that we can make $Q = 0$ for values of $\lambda_1, \cdots, \lambda_6$ in F not all zero. Define d by (47) and f by (48) for the particular λ_i we have used to make Q vanish. Since A is the *direct* product of B and C, the quantities $d - f$ and $d + f$ are both not zero when the λ_i are not all zero. But $(d-f)(d+f) = d^2 - f^2 = Q_1 - Q_2 = Q = 0$. Hence in A a product of two non-zero quantities is zero and A is not a division algebra.

Conversely, let Q not be a null form. Then, in particular, Q_1 and Q_2 are not null forms and B and C are known* to be division algebras. The algebra Γ whose quantities have the form $X = x_1 + x_2 u$, where x_1 and x_2 are in C, has a division sub-algebra C and the property that if we define $x' = x$ for every x of C, then $u^2 = \tau$ in C, $x'' = (x')' = u^2 x u^{-2} = x$ for every x of C. But then Γ is known† to be a division algebra if and only if $\tau \neq x'x = x^2$ for any x of C. But τ is in F and if $\tau = x^2$ then, since x is an f of (48), and u is a d of (47), we have $Q = 0$ for $\lambda_1 = 1$, a contradiction of our hypothesis that Q was not a null form.

Define $X' = x_1 - x_2 u$, for every X of Γ, and we will have $X' = sXs^{-1}$, $X'' = s^2 X s^{-2} = X$, $s^2 = \sigma$ in F. Then it is known

* See L. E. Dickson, *Algebren und ihre Zahlentheorie*, p. 47, for the condition $\sigma \neq \xi_1^2 - \xi_2^2\tau$, equivalent to the condition we have stated.

† A theorem of L. E. Dickson, ibid., pp. 63–64.

(Dickson, loc. cit.) that A, whose quantities have the form $X+Ys$, is a division algebra when Γ is one if and only if

$$s^2 = X'X \text{ for any } X \text{ of } \Gamma.$$

But if $s^2 = X'X$, $(x_1 - x_2u)(x_1 + x_2u) = x_1^2 - x_2^2\tau + (x_1x_2 - x_2x_1)u = \sigma$ we have

$$\sigma = x_1^2 - x_2^2\tau, \quad x_1x_2 = x_2x_1. \tag{49}$$

First let x_1 and x_2 be in F. Then Q is a null form when we take $\lambda_1 = \sigma$, $\lambda_2 = \tau x_2$, $\lambda_3 = x_1$, $\lambda_4 = \lambda_5 = \lambda_6 = 0$, since $0 = \sigma\tau(\sigma + x_2^2\tau - x_1^2) = \sigma^2\tau + (\tau x_2)^2\sigma - x_1^2\sigma\tau$, a contradiction. Next let x_1 be in F but x_2 not in F. Then $x_2^2\tau = x_1^2 - \sigma$ is in F and $x_2\tau$ is an f of (48) while $(x_2\tau)^2 = Q_2 = x_1^2\tau - \sigma\tau$ so that Q is a null form for $Q_2 = (x_2\tau)^2$, $\lambda_1 = x_1$, $\lambda_2 = 0$, $\lambda_3 = 1$. The only remaining case is where x_1 is not in F. If x_1^2 were in F so that x_1 would be an f of (48), then $x_1x_2 = x_2x_1$ implies that $x_2 = \xi + \eta x_1$, ξ and η in F, since in a generalized quaternion division algebra the only quantities commutative with a non-scalar quantity x are scalar coefficient polynomials in x. But x_2^2 is in F so that $\eta = 0$ or $\xi = 0$. When $\eta = 0$, then $x_1^2 = \xi^2\tau + \sigma$ and $Q_2 = Q_1 = \xi^2\tau + 1^2\sigma - O\sigma\tau$, a contradiction of our hypothesis. When $\xi = 0$ then $x_2 = \eta x_1$ and $x_1^2 - x_2^2\tau = x_1^2(1 - \eta^2\tau)$. But by Lemma 1 we have $(1-\eta^2\tau)^{-1} = \delta_1^2 - \delta_2^2\tau$ and $x_1^2 = Q_2 = \sigma\delta_1^2 - \sigma\tau\delta_2^2 = Q_1$, a contradiction. We have finally come to the case where neither x_1 nor its square is in F. We then have, where f is given by (47) and $f^2 = Q_2$, that $x_1 = \lambda_7 + f$ with $\lambda_7 \neq 0$ in F. As before the relation $x_1x_2 = x_2x_1$ implies that x_2 is a polynomial in x_1. But now we may write $x_2 = \xi + \eta f$. Now

$$x_1^2 - x_2^2\tau = \lambda_7^2 + 2\lambda_7 f + Q_2 - (\xi^2 + 2\xi\eta f + \eta^2 Q_2)\tau = \sigma.$$

It follows that $2\lambda_7 - 2\xi\eta\tau = 0$, so that $\lambda_7 = \xi\eta\tau$ and

$$\sigma = \xi^2\eta^2\tau^2 - \xi^2\tau + Q_2(1 - \eta^2\tau) = (Q_2 - \xi^2\tau)(1 - \eta^2\tau).$$

The quantity $(1 - \eta^2\tau) \neq 0$ has an inverse $\delta_1^2 - \delta_2^2\tau$ with δ_1 and δ_2 in F by Lemma 1, and $Q_2 - \xi^2\tau = \sigma(\delta_1^2 - \delta_2^2\tau)$, so that we have $Q_2 = \xi^2\tau + \sigma\delta_1^2 - \sigma\tau\delta_2^2 = Q_1$. We have again shown that if A were not a division algebra, then Q would be a null form, a contradiction of our hypothesis. Hence A is a division algebra and we have proved the following theorem

THEOREM 3. *A direct product $A=B\times C$ of two generalized quaternion algebras $B=(1, u, s, us)$, $C=(1, j, t, jt)$ with $u^2=\tau$, $s^2=\sigma$, $su=-us$, $j^2=\gamma$, $t^2=\rho$, $tj=-jt$, is a division algebra if and only if the quadratic form*

$$(50)\qquad Q=(\lambda_1^2\tau+\lambda_2^2\sigma-\lambda_3^2\sigma\tau)-(\lambda_4^2\gamma+\lambda_5^2\rho-\lambda_6^2\gamma\rho)$$

in the variables $\lambda_1, \lambda_2, \cdots, \lambda_6$ in F, is not a null form.

We may now apply Theorem 3 and our previous results to obtain complete necessary and sufficient conditions that a cyclic algebra be a division algebra. We first assume that $\gamma^2=N(a)$ for some a in $F(i)$. If $\gamma=0$, then $-\gamma=\beta_1^2-\beta_2^2\tau$ with $\beta_1=\beta_2=0$, and the form Q may be defined. If $\gamma\neq 0$, then by Corollary 2 we can again define the form Q with

$$\sigma=2\beta_1\gamma, \rho=2\nu\tau\beta_1(\beta_2\Delta-\beta_1).$$

Suppose first that Q is a null form. If $\gamma=0$, then $y^4=0$ while y is not zero and A is not a division algebra. If $\gamma\neq 0$ but $\beta_1=0$ or $\beta_2\Delta-\beta_1=0$, then again, as we have seen, A is not a division algebra. The only other case is where Theorem 2 can be applied and, by Theorem 3, A is again not a division algebra. Conversely let Q be not a null form. Then obviously from our definition of Q as above and the fact that we have the coefficients of Q all not zero in a non-null form, $\gamma\neq 0$, $\beta_2\Delta\neq\beta_1$, $\beta_1\neq 0$, and again A is the direct product of B and C; we may again apply Theorem 3, and A is a division algebra.

THEOREM 4. *Let A be a cyclic algebra with basis $i^\lambda y^\mu$, $(\lambda, \mu=0, 1, 2, 3)$, where i is a root of the cyclic quartic*

$$\phi(\omega)\equiv\omega^4+2\nu\tau\omega^2+\nu^2\Delta^2\tau=0$$

with $\tau=1+\Delta^2$, $\nu\neq 0$, $\Delta\neq 0$ in F, and τ not the square of any element in F. Also

$$\theta(i)=qi, q\Delta=1+u, i^2=\nu(u-\tau), yi=\theta(i)y, y^4=\gamma \text{ in } F.$$

Suppose that γ^2 is the norm of a quantity of $F(i)$ so that we have $-\gamma=\beta_1^2-\beta_2^2\tau$ with β_1 and β_2 in F. Then A is a division algebra if and only if the form

$$Q=\lambda_1^2\tau+\lambda_2^2\sigma-\lambda_3^2\sigma\tau-\lambda_4^2\gamma-\lambda_5^2\rho+\lambda_6^2\gamma\rho$$

with $\sigma=2\beta_1\gamma$, $\rho=2\beta_1\nu\tau(\beta_2\Delta-\beta_1)$ does not vanish for any $\lambda_1, \cdots, \lambda_6$ not all zero and in F.

The only other case is $\gamma^2 \neq N(a)$. Then obviously $\gamma \neq N(a)$ since otherwise $\gamma^2 = N(a^2)$, a contradiction. If $\gamma^3 = N(a)$, then either $\gamma = 0$, whence $\gamma^2 = N(0)$, a contradiction, or else $\gamma \neq 0$, $\gamma^6 = \gamma^2\gamma^4 = N(a^2)$, $\gamma^2 = N(a\gamma^{-1})$, again a contradiction. Hence the condition $\gamma^2 \neq N(a)$ is equivalent to the Wedderburn norm condition. We have also shown the former condition equivalent to the condition $\gamma \neq -ee'$ for any e of $F(u)$. We thus have proved the following theorem

THEOREM 5. *Let all the hypotheses of Theorem* 4 *be satisfied except that now* $\gamma^2 \neq N(a)$ *for any* a *of* $F(i)$, *or, what is the same thing,* $-\gamma$ *is not expressible in the form* $\beta_1^2 - \beta_2^2\tau$, β_1 *and* β_2 *in* F. *Then the cyclic algebra* A *is a division algebra.*

We shall finally pass to the case where F is the field R of all rational numbers. Quadratic forms have been studied in detail for this case and it has been shown that every indefinite quadratic form in five or more variables is a null form.* The numbers τ, σ, $-\sigma\tau$ all have the same sign only when all are negative. If they are all negative and γ, ρ, $-\gamma\rho$ are also all negative then τ and $-\gamma$ have opposite signs so that $Q = Q_1 - Q_2$ is an indefinite quadratic form. In the other cases obviously Q is indefinite, providing that its coefficients are all not zero. When some of the coefficients of Q are zero then, by making all the other variables zero and those with zero coefficients not zero, we can make Q zero so that Q is a null form. When none of the coefficients of Q is zero then Q is an indefinite quadratic form in six variables and hence is a null form. Hence in every case the cyclic algebra A is not a division algebra when the hypotheses of Theorem 4 are satisfied. We have† the following result.

THEOREM 6. *When* $F = R$, *the field of all rational numbers, the Wedderburn norm condition for cyclic algebras of order sixteen is necessary as well as sufficient.*

COLUMBIA UNIVERSITY

* For the first complete proof of this theorem see L. E. Dickson, *Studies in the Theory of Numbers.*

† We also have here a new short proof of the author's theorem that a direct product of two rational generalized quaternion division algebras is never a division algebra, by using the above proof that *when* Q *is a null form* A is not a division algebra. This theorem was first proved by the author and published in the Annals of Mathematics, vol. 30 (1929), pp. 621–625.

OFFICIAL COMMUNICATIONS

Meetings of the Society have been fixed at the following times and places:

SEATTLE, WASHINGTON, June 13, 1931.

Abstracts must be in the hands of Associate Secretary T. M. Putnam, University of California, Berkeley, Calif., not later than May 16. For publication in advance of the meeting, abstracts must be received not later than May 9.

MINNEAPOLIS, MINNESOTA, SUMMER MEETING AND COLLOQUIUM, September 8-11, 1931.

Abstracts must be in the hands of Associate Secretary M. H. Ingraham, University of Wisconsin, Madison, Wis., not later than August 10. For publication in advance of the meeting, abstracts must be received not later than July 7. Professor Marston Morse will deliver a series of colloquium lectures on *The calculus of variations*. Professor C. C. MacDuffee will deliver an address entitled *Ideals in the theory of linear algebras*.

R. G. D. RICHARDSON, *Secretary of the Society.*

Articles for insertion in the BULLETIN should be addressed to E. R. HEDRICK, Editor of the BULLETIN, University of California at Los Angeles. Reviews should be sent to W. R. LONGLEY, Yale University New Haven, Conn. Notes should be sent to H. W. KUHN, Ohio State University, Columbus, Ohio.

Subscriptions to the BULLETIN, orders for back numbers, and inquiries in regard to non-delivery of current numbers should be addressed to the American Mathematical Society, 450-459 Ahnaip St., Menasha, Wis., or 501 West 116th St., New York.

Advertising space is available in the BULLETIN at $16 per page, $9 per half page, $5 per quarter page, per issue. Address correspondence regarding advertising to H. L. Rietz, University of Iowa, Iowa City, Iowa.

Changes of address of members should be communicated at once to the Secretary of the Society, R. G. D. Richardson, 501 West 116th St., New York.

The initiation fees and the annual dues of members of the Society (see this BULLETIN, p. 322, May, 1930; and the List of Officers and Members, October, 1930, p. 58), are payable to the Treasurer of the Society, G. W. Mullins, 501 West 116th St., New York.

WHOLE NUMBER 379

CONTENTS

	PAGE
Linear Functional Transformations in General Spaces. By T. H. HILDEBRANDT	185
On the Zeros of Exponential Sums and Integrals. By R. E. LANGER	213
Types of Series and Types of Summability. By C. N. MOORE	240
Factorability of Numerical Functions. By E. T. BELL	251
Two Types of Connected Sets. By P. M. SWINGLE	254
Quadrilaterals Inscribed and Circumscribed to a Plane Cubic. By A. H. DIAMOND	258
On a Type of Illusory Theorem concerning Higher Indeterminate Equations. By E. T. BELL	261
Parametric Solutions of certain Diophantine Equations. By T. A. PIERCE	264
Sums of Four or More Values of $\mu x^2+\nu x$ for Integers x. By GORDON PALL	267
Groups Generated by Two Operators whose Squares are Invariant. By G. A. MILLER	270
On a Problem of N. Aronszajn and an Axiom of R. L. Moore. By LEO ZIPPIN	276
Statistics of a Set of Closed Intervals. By F. H. MURRAY	281
Extension of a Theorem of Mazurkiewicz. By R. L. WILDER	287
A Property of Continua Similar to Local Connectivity. By W. A. WILSON	294
On the Wedderburn Norm Condition for Cyclic Algebras. By A. A. ALBERT	301

For official Communications and Notices, see the inside of the back cover.